

TRANSACTIONS OF  
THE ROYAL SOCIETY  
OF CANADA

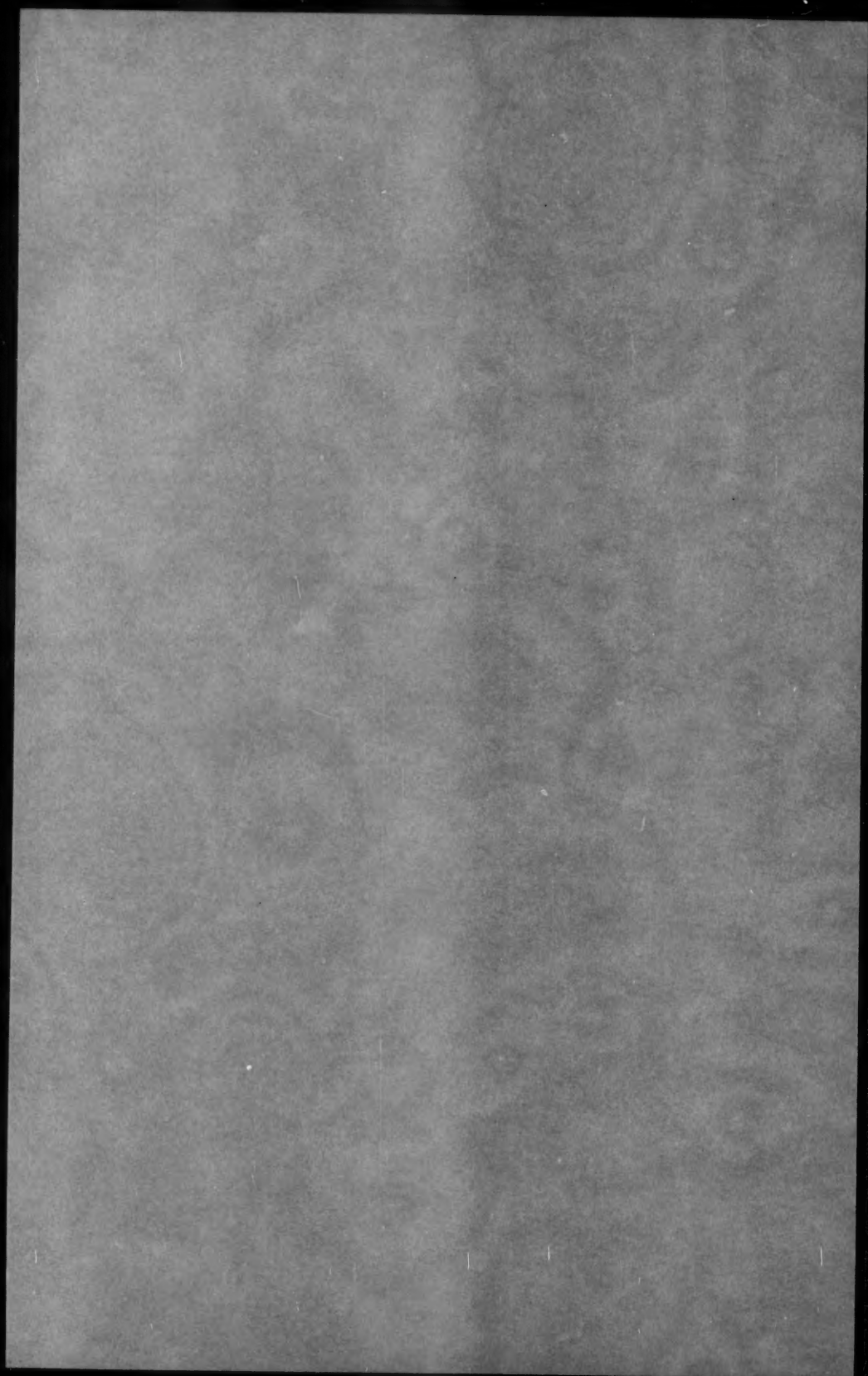
SECTION III  
CHEMICAL, MATHEMATICAL,  
AND  
PHYSICAL SCIENCES



THIRD SERIES—VOLUME LI—SECTION III

JUNE, 1957

OTTAWA  
THE ROYAL SOCIETY OF CANADA  
1957



TRANSACTIONS OF  
THE ROYAL SOCIETY  
OF CANADA

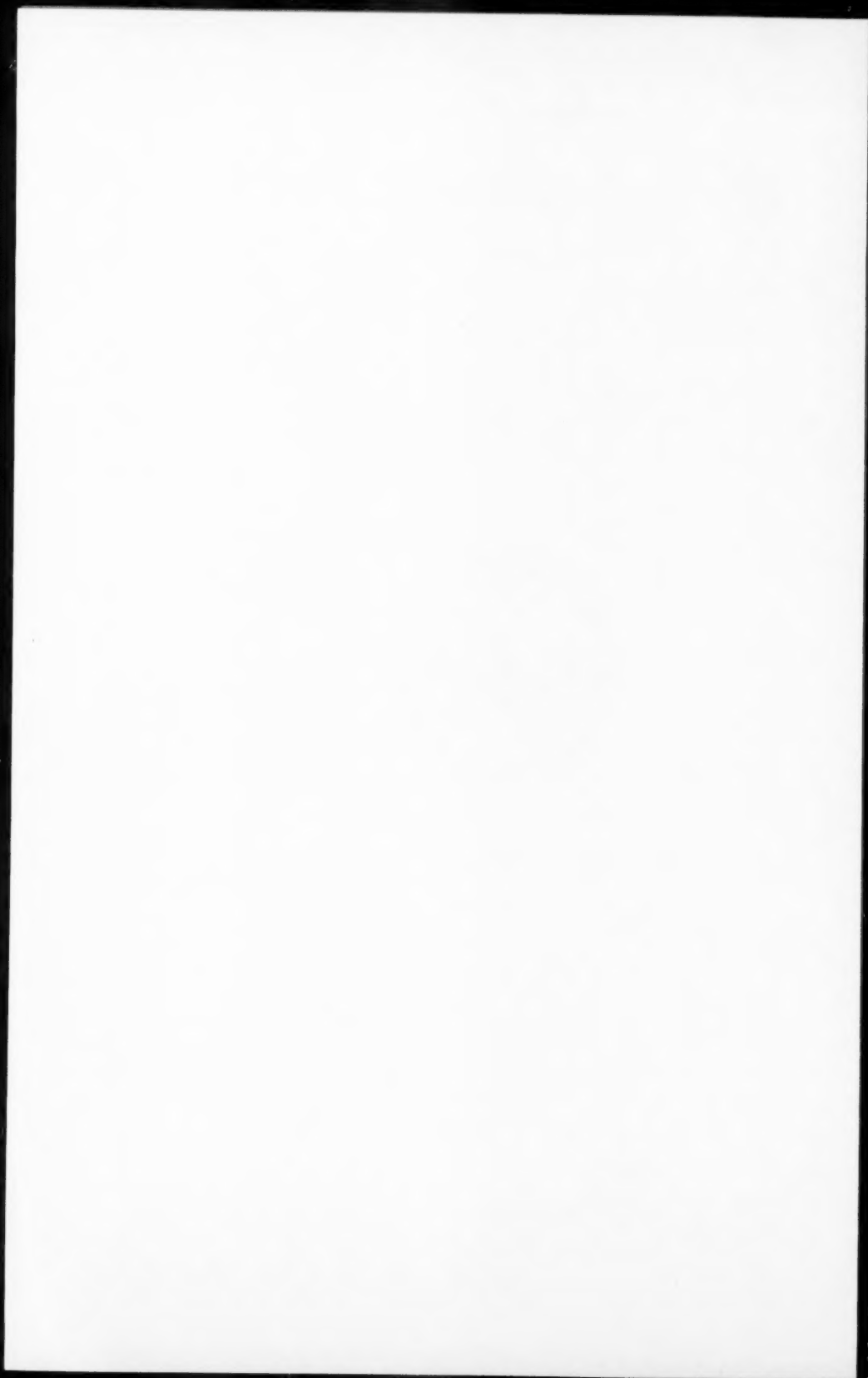
SECTION III  
CHEMICAL, MATHEMATICAL,  
AND  
PHYSICAL SCIENCES



THIRD SERIES—VOLUME LI—SECTION III

JUNE, 1957

OTTAWA  
THE ROYAL SOCIETY OF CANADA  
1957



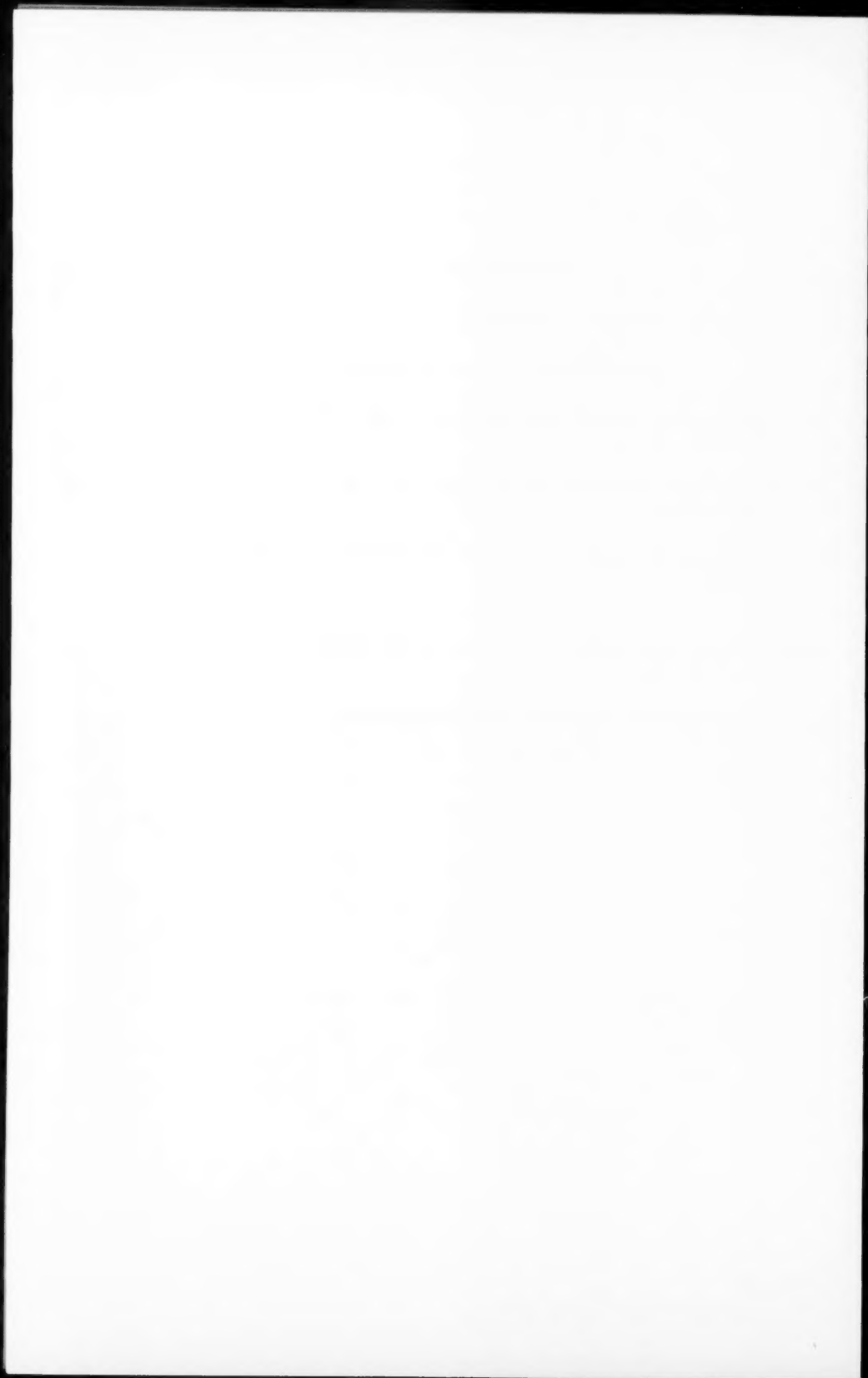


## CONTENTS



### A SYMPOSIUM ON SYMMETRY

<i>I. Crystal Symmetry and Its Generalizations.</i> By H. S. M. COXETER, F.R.S.C. . . . .	1
<i>II. Symmetry and Interaction between Elementary Particles.</i> By W. OPECHOWSKI . . . . .	15
<i>III. The Symmetry Sense in Chemistry.</i> By GEORGE F WRIGHT, F.R.S.C. . . . .	31
<i>Differentiability Properties of Arcs of Order <math>n + 1</math> in Conformal <math>n</math>-space.</i> By N. D. LANE . . . . .	45
<i>A Theorem of Friedrichs.</i> By HANS ZASSENHAUS, F.R.S.C. . . . .	55



## A SYMPOSIUM ON SYMMETRY

I. Crystal Symmetry and Its Generalizations<sup>1</sup>

H. S. M. COXETER, F.R.S.C.

The topic for this symposium is the general idea of symmetry in some of its many aspects. According to the Oxford Dictionary, *symmetry* means "divisibility into two or more parts, each of the same shape and size as the others and similarly placed with respect to the dividing points or lines or planes." This description comes very close to the mathematical definition: the possession of a group of *automorphisms*, usually congruent transformations (namely reflections, rotations, translations, and the combinations of these in pairs). For instance, the quadrirectangular tetrahedron or *orthoscheme*

$$(-1, -1, -1) (-1, -1, 1) (-1, 1, 1) (1, 1, 1)$$

is symmetrical by the transformation

$$(x, y, z) \rightarrow (-z, -y, -x),$$

which is a half-turn (i.e., rotation through two right angles) about the line

$$x + z = 0, \quad y = 0.$$

Among such transformations, the fundamental one, of which all the others are combinations, is *reflection*. An orthoscheme and its mirror image are congruent (in the sense that corresponding distances within them are equal) but are not superposable by any motion. They are related like right- and left-handed screws. In the terminology of Leibniz, they are *indiscernible*. According to the late Professor Hermann Weyl (10, p. 17), "the inner structure of space does not permit us, except by arbitrary choice, to distinguish a left from a right screw. [On] this fundamental notion . . . depends the entire theory of relativity, which is but another aspect of symmetry."

Any object, however irregular, becomes symmetrical when we place it beside its image in a mirror. This simplest kind of symmetry, *bilateral* symmetry, is characteristic of the external shape of all animals more highly organized than the lobster. It received its supreme poetic expression in the words of William Blake:

<sup>1</sup>Presidential Address to Section III, June 1957.

Tyger! Tyger! burning bright  
 In the forests of the night,  
 What immortal hand or eye  
 Dare frame thy fearful symmetry?

The impressiveness of symmetry is well illustrated by a remark of Sir D'Arcy Wentworth Thompson (in a letter of March 1947): "The Great Pyramid surpasses all expectation, and beggars all description. I remember being somewhat disappointed with Niagara! But the Great Pyramid never disappointed anybody."

Of course, the Pyramid is far more symmetrical than the tiger. The symmetry group of the tiger is of order 2, generated by a single reflection. That of the square pyramid is of order 8, generated by two reflections, in mirrors inclined at  $45^\circ$  so as to form a sort of kaleidoscope. The orthoscheme mentioned above is one-eighth of a square pyramid. When we place it between the mirrors, we see the whole pyramid. (Figure 1 shows this in "plan.")

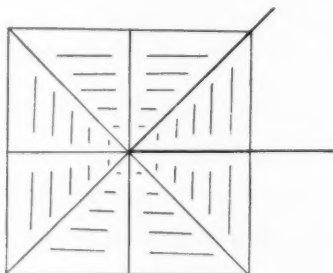


FIGURE 1

A third mirror, in the position of the sloping face, would exhibit this pyramid as one of six, fitting together to make a cube. Since each pyramid is composed of 8 orthoschemes, the whole cube is composed of 48; in fact, the symmetry group of the cube is of order 48.

Finally, we could add a fourth mirror in the position of the remaining face of the orthoscheme, which is the base of the pyramid, that is, a face of the cube. Then the cube is reflected into an infinite honeycomb of cubes filling all space: a pattern whose symmetry group is infinite (2, p. 71). The same cubic honeycomb could have been produced by six mirrors instead of four, namely, one in each face of a cube, as if we stood in a square room with a mirror in each wall as well as in the ceiling and floor. Theoretically, we would see infinitely many images of ourselves, some standing upright, others inverted.

Such patterns occur in nature as the positions of atoms in a crystal. For instance, if we imagine the cubes of the cubic honeycomb to be coloured alternately black and white, like a three-dimensional chess-board, we

obtain the arrangement of atoms in a crystal of common salt, with a sodium atom in each black cube and a chlorine atom in each white cube. If we dissect each white cube into six pyramids and attach each pyramid to the neighbouring black cube (2, p. 26), we obtain a honeycomb of rhombic dodecahedra (a shape that occurs in nature as a crystal of garnet). Thus the symmetry group of the honeycomb of rhombic dodecahedra is a subgroup of index 2 in the symmetry group of the honeycomb of cubes.

The enumeration of such *space groups* is the central problem of mathematical crystallography: a very complicated problem which was solved about 70 years ago by Fedorov in Russia, Schoenflies in Germany, and Barlow in England. All three found independently that there are exactly 230 distinct groups. To give some idea of the nature of such an infinite symmetry group, let us reduce the number of dimensions from three to one (or perhaps I should say to  $1\frac{1}{2}$ ) and consider the seven ways of repeating a pattern on a strip or ribbon (9, pp. 81-82).

Typical pattern	Generators	Abstract group
b b b b . . .	1 translation	} $\mathbb{C}_\infty$
b p b p . . .	1 glide-reflection	
V V V V . . .	2 reflections	} $\mathbb{D}_\infty$
N N N N . . .	2 half-turns	
V A V A . . .	1 reflection and 1 half-turn	$\mathbb{C}_\infty \times \mathbb{C}_2$
D D D D . . .	1 reflection and 1 translation	
H H H H . . .	3 reflections	$\mathbb{D}_\infty \times \mathbb{C}_2$

In the third and fifth groups, the generating reflections are in vertical lines. In the sixth group the reflection is in a horizontal line, and of course the translation is in a horizontal direction. The product of this reflection and translation is the so-called *glide-reflection*, which generates the second group.

The first two groups are isomorphic. Abstractly, either is  $\mathbb{C}_\infty$ , the free group with one generator. The next three are again isomorphic: the two generators, say  $P$  and  $Q$ , satisfy the relations

$$P^2 = Q^2 = 1.$$

This group is denoted by  $\mathbb{D}_\infty$ . The last two groups are direct products of  $\mathbb{C}_\infty$  and  $\mathbb{D}_\infty$  with the group of order 2 generated by the horizontal reflection.

In saying that these are patterns in  $1\frac{1}{2}$  dimensions, I mean that they are in a plane but involve translation in only one direction. From the standpoint of a purely one-dimensional creature, the horizontal reflection has no effect. For him there are only two space-groups:  $\mathbb{C}_\infty$ , generated by a translation, and  $\mathbb{D}_\infty$ , generated by two reflections (whose product is the translation). These symbols are used because the one-dimensional space-groups are limiting cases of the finite groups of congruent transformations in two dimensions: the *cyclic* group  $\mathbb{C}_n$  generated by a rotation through one- $n$ th of a whole turn, and the *dihedral* group  $\mathbb{D}_n$  generated by reflections in two



PLATE I

mirrors inclined at half this angle. For instance,  $\mathcal{C}_4$  is the group of the swastika, and  $\mathcal{D}_4$  is the group of the square, or of the square pyramid. Similarly,  $\mathcal{D}_6$  is the group of the regular hexagon, or of the snowflake (1).

It is natural to consider next the two-dimensional patterns with translations in more than one direction. Fricke and Klein (6, pp. 227-233) showed that there are just 17 such 2-dimensional space-groups. A few of the simplest can be seen in any ordinary wallpaper. One of the most complicated (in the form of a tessellation of regular hexagons) was invented millions of years ago by the bees (see Plate I). In this case the group is generated by reflections in the sides of a triangle with angles  $30^\circ$ ,  $60^\circ$ ,  $90^\circ$ . The symmetry of each hexagon is the same as that of the snowflake, but now there is a third reflection interchanging two adjacent hexagons.

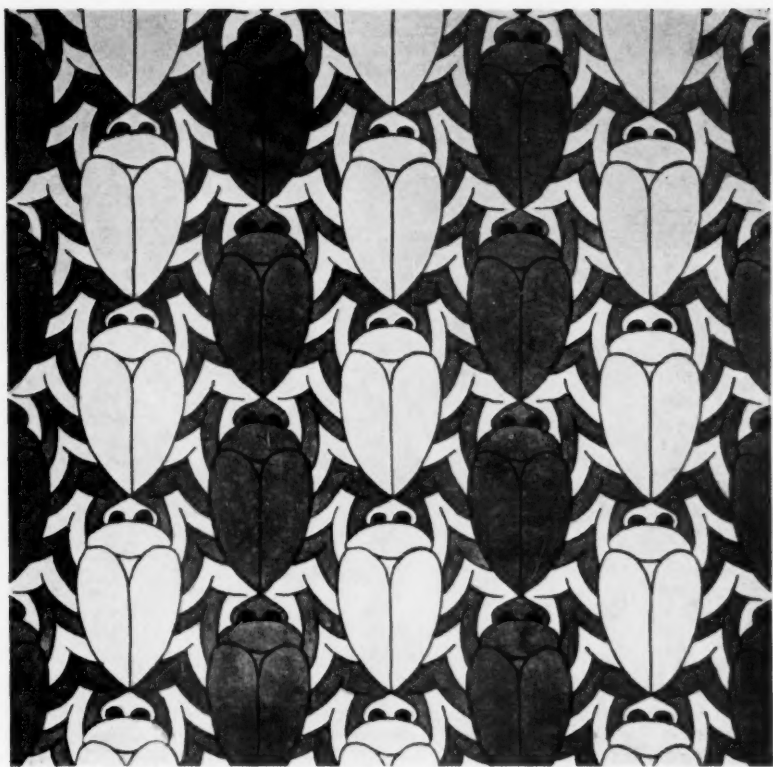


PLATE II

As an art form, the making of plane patterns reached its highest development in thirteenth-century Spain, where the Moors unconsciously used all the seventeen groups in their intricate decoration of the Alhambra (7). Their preference for abstract patterns was due to their strict observance of the Second Commandment. In our own time, the Dutch artist M. C. Escher, free from such scruples, illustrates some of these groups by using animal shapes for their fundamental regions. For instance, the group of his pattern of beetles (Plate II) seems at first sight to be the group **pm** generated by two vertical reflections and a vertical translation (4, pp. 42-44). But on looking more closely we see that there are both dark and light beetles, and these are interchanged by a glide-reflection. The whole group **cm** is generated by this vertical glide-reflection and a vertical reflection. By repeating ("squaring") the glide-reflection, we obtain the translation mentioned above.

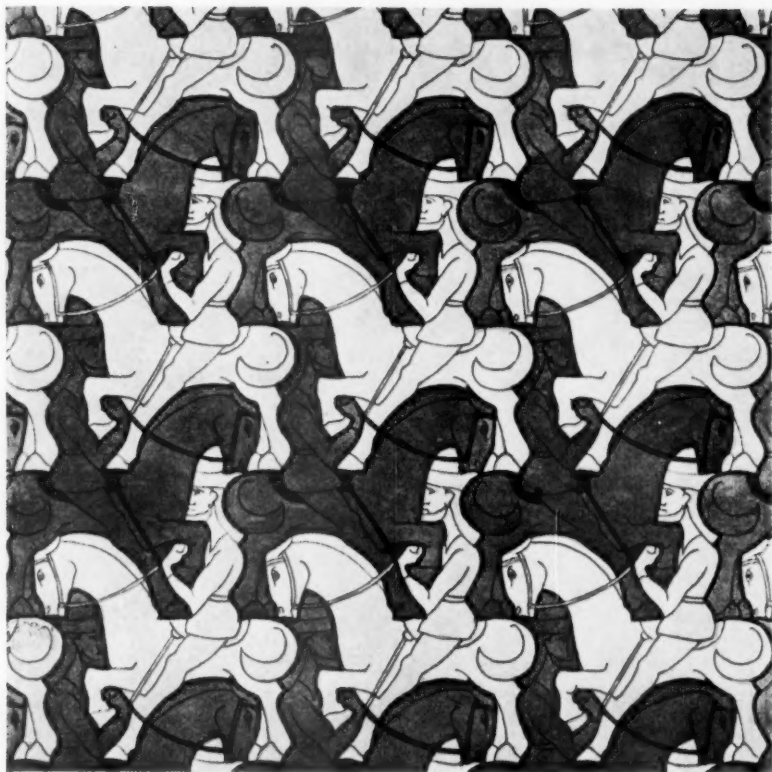


PLATE III

Similarly, the group of Escher's pattern of knights on horseback (Plate III) seems at first sight to be **p1**, generated by two translations (4, pp. 40, 43). But by ignoring the distinction between the dark and light specimens we obtain an interesting group **pg**, generated by two parallel glide-reflections, say  $P$  and  $Q$ . We observe that the vertical translation can be expressed equally well as  $P^2$  or  $Q^2$ . It is remarkable that the equation

$$P^2 = Q^2$$

constitutes a complete abstract definition for the group. This means that every relation satisfied by  $P$  and  $Q$  is an algebraic consequence of this single relation.

We come still closer to crystallography by considering the *point groups*: finite groups of congruent transformations (including the 32 "crystal classes"). Given such a group and an arbitrary point, consider all the



transforms of the point. This finite set of points is transformed into itself by every element of the group. Hence there is an invariant point: the centroid (or centre of gravity) of all the points. Taking this as origin, we have a finite group of *orthogonal* transformations. Each transformation may be specified by its effect on a frame of three coordinate axes. It follows (2, p. 36) that every orthogonal transformation is either a reflection or the product of two or three reflections, and that every *direct* transformation (preserving sense) is the product of just two reflections, that is, a rotation. Incidentally, this is the easiest way to prove that the product of two rotations (about lines through the origin) is another rotation.

The finite groups of rotations (2, pp. 53–55) are found to be the cyclic group  $\mathbb{C}_n$ , the dihedral groups  $\mathbb{D}_n$  (generated by "reflections in lines", i.e., by half-turns) and the rotation groups of the Platonic solids. The remaining groups of orthogonal transformations (10, p. 155) are easily derived by combining the rotation groups with a very special sense-reversing transformation: the *central inversion* (or "reflection in a point"). This is the product of reflections in three mutually perpendicular mirrors. By looking at yourself between two perpendicular mirrors you see yourself as others see you. But the third mirror, like the first, reverses sense; it also turns you upside-down. When the three mirrors are taken as coordinate planes, each reflection reverses the sign of one coordinate, and the three together reverse the signs of all three.

On a sphere centred at the origin, the central inversion interchanges every pair of antipodal points. By abstractly identifying such pairs of points, we pass from spherical geometry to the *elliptic* geometry of Cayley and Klein, in which every two "lines" have a unique point of intersection. This abstract identification was vividly described by H. G. Wells in his short story *The Remarkable Case of Davidson's Eyes*. Because of some catastrophe, Davidson's field of vision was so distorted that he saw everything as it would have appeared from an exactly antipodal position on the Earth.

Riemann and Einstein suggested that our astronomical space might be finite but unbounded. In this case it might be either *spherical* 3-space, like the hypersurface of a hypersphere in four dimensions, or *elliptic* 3-space, derived by identifying antipodes. The four-dimensional central inversion, being the product of four reflections, does not reverse sense. In the words of the late Sir Arthur Eddington (5, p. 158): "We may leave to the metaphysicist the question whether two objects can be exactly alike, both intrinsically and in relation to all surroundings, and yet differ in identity."

Returning to finite groups of orthogonal transformations, I would like to show you some objects possessing symmetry groups of this type. The complete group of the regular tetrahedron is of order 24, since it is the symmetric group consisting of the  $4!$  permutations of the four vertices. (For instance, any two vertices can be transposed by a reflection.) Another representation of the same group is provided by the rotational symmetry-

operations of the cube. Each rotation is expressible as a permutation of the four "diameters" joining pairs of opposite vertices of the cube. An interesting figure having such a symmetry group is the *snub cube* (Fig. 2), one of the thirteen Archimedean solids (3, p. 439 (fig. 24)).

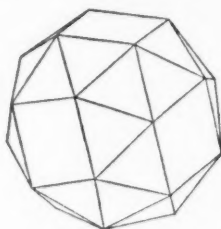


FIGURE 2

These twenty-four permutations fall into two sets of twelve, called *even* and *odd* permutations. The distinction is illustrated by Kepler's *stella octangula* (Fig. 3), which consists of two regular tetrahedra whose edges are the diagonals of the faces of a cube. The even permutations of the four diameters (which form the *alternating* group of order 12) rotate each tetrahedron into itself; the odd permutations interchange the two tetrahedra.

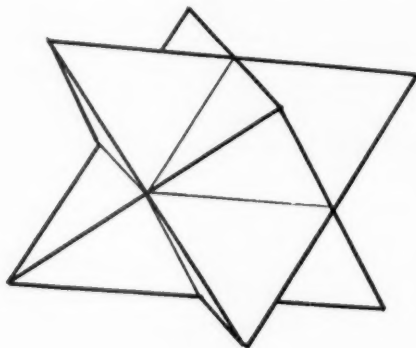


FIGURE 3

Closely resembling these crystallographic point groups are the non-crystallographic point groups, such as the symmetry group of the regular dodecahedron, and its rotational subgroup. This subgroup is the alternating group of order 60, consisting of all the even permutations of five objects. To see this, we take the five objects to be five regular tetrahedra inscribed in the dodecahedron (Fig. 4).<sup>2</sup> Each rotation corresponds to an even permutation of them: a double transposition, a cyclic permutation of three, or a

<sup>2</sup>Figures 4, 5 and 6 were drawn by J. Flinders Petrie.



FIGURE 4

cyclic permutation of all five. The last is a rotation of period five, which cannot occur as a symmetry-operation of a crystal, though of course it can occur among living things such as a five-petalled flower or a starfish.

Figure 5 is a picture of a sphere divided into spherical triangles (24 black and 24 white) by nine great circles lying in the planes of symmetry of a

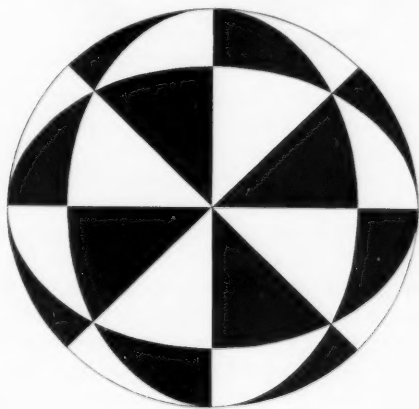


FIGURE 5

cube: three planes parallel to the faces, and six joining pairs of opposite edges. The vertices of the cube itself appear as the points where the angles are  $60^\circ$  so that three triangles of each colour come together. We observe that each triangle (of either colour) is transformed into its neighbours (of the other colour) by reflections in its sides. These three reflections generate the extended octahedral group of order 48: the complete symmetry group of the cube. Its rotational subgroup consists of those transformations which take black triangles into black triangles, and white into white. Thus the 24 triangles of either colour represent the  $4!$  permutations of the four diameters of the cube.

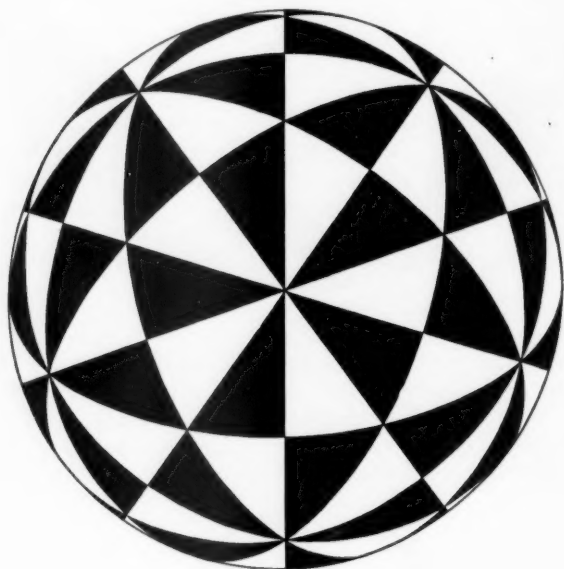


FIGURE 6

Figure 6 shows analogously the sphere divided into 60 black and 60 white triangles by the 15 planes of symmetry of the regular dodecahedron. Our familiarity with three-dimensional space enables us to accept the idea that these triangles are all the same size even though the peripheral ones are made to look smaller by perspective foreshortening. Instead of orthogonal projection, we might have used stereographic projection, spreading out the sphere into the whole plane (8). Then we would have a system of fifteen circles, meeting at the proper angles, and the reflections would be replaced by inversions. This means that we would have a group generated by inversions in three circles.

In Figure 7 we see another such group, with the important difference that now the angle-sum of each triangle is less than two right angles and the number of triangles is infinite. The group is again generated by inversions in three circles, but the figure is no longer a picture of something in space. We do not find it as easy as before to imagine that the smaller peripheral triangles are the same size as those in the middle. But in so far as we succeed in stretching our imagination to this extent, we are visualizing the non-Euclidean plane of Gauss, Bolyai and Lobatschewsky.

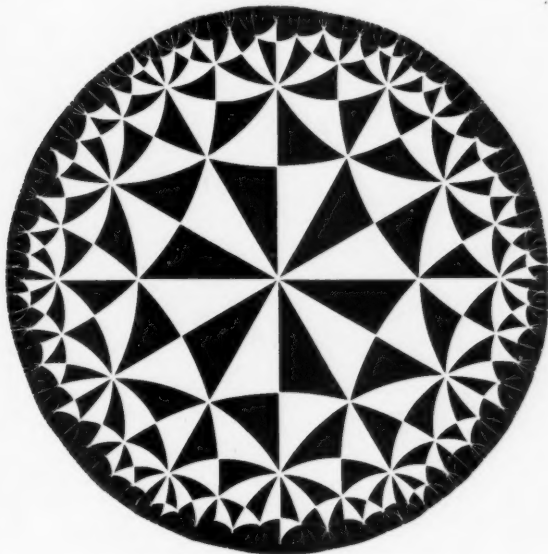


FIGURE 7

This is one way to generalize the idea of symmetry. Another is to increase the number of dimensions. Plate IV shows a wire model made by Mr. P. S. Donchian of Hartford, Conn.<sup>3</sup> This represents an orthogonal projection of a four-dimensional hyper-solid bounded by 120 regular dodecahedra. The model has the same 120 symmetry-operations as a single dodecahedron, but the four-dimensional polytope itself has a symmetry group of order  $120^2 = 14400$ .

Yet another generalization is from real space to complex space, where the period of a reflection may be greater than 2, so that instead of the customary object and image we have an object and several images, all in a single mirror! For instance, two squares, each inscribed in the other,

<sup>3</sup>For other views of the same model, see (2, Plates V and VIII).

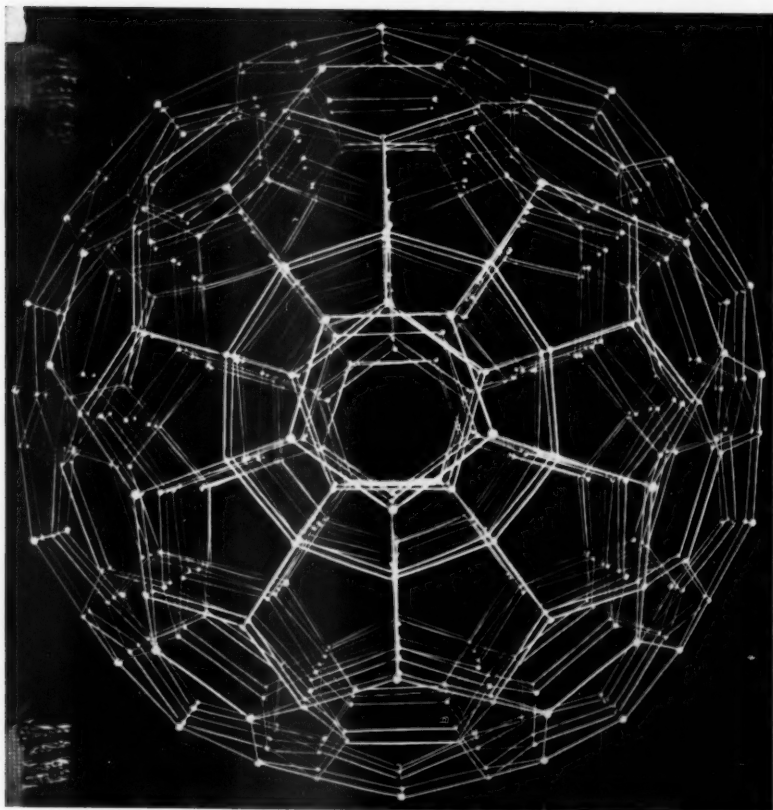


PLATE IV

cannot be constructed in the real plane; but in the unitary plane they can form a regular complex polygon which is transformed into itself by a group of order 24, generated by two reflections of period 3. This group (4, p. 76) is abstractly defined by the two relations

$$P^3 = 1, \quad PQP = QPQ$$

(which imply  $Q^3 = 1$ ), whereas the group  $\mathfrak{D}_3$  of order 6, generated by two ordinary mirrors inclined at  $60^\circ$ , is defined by

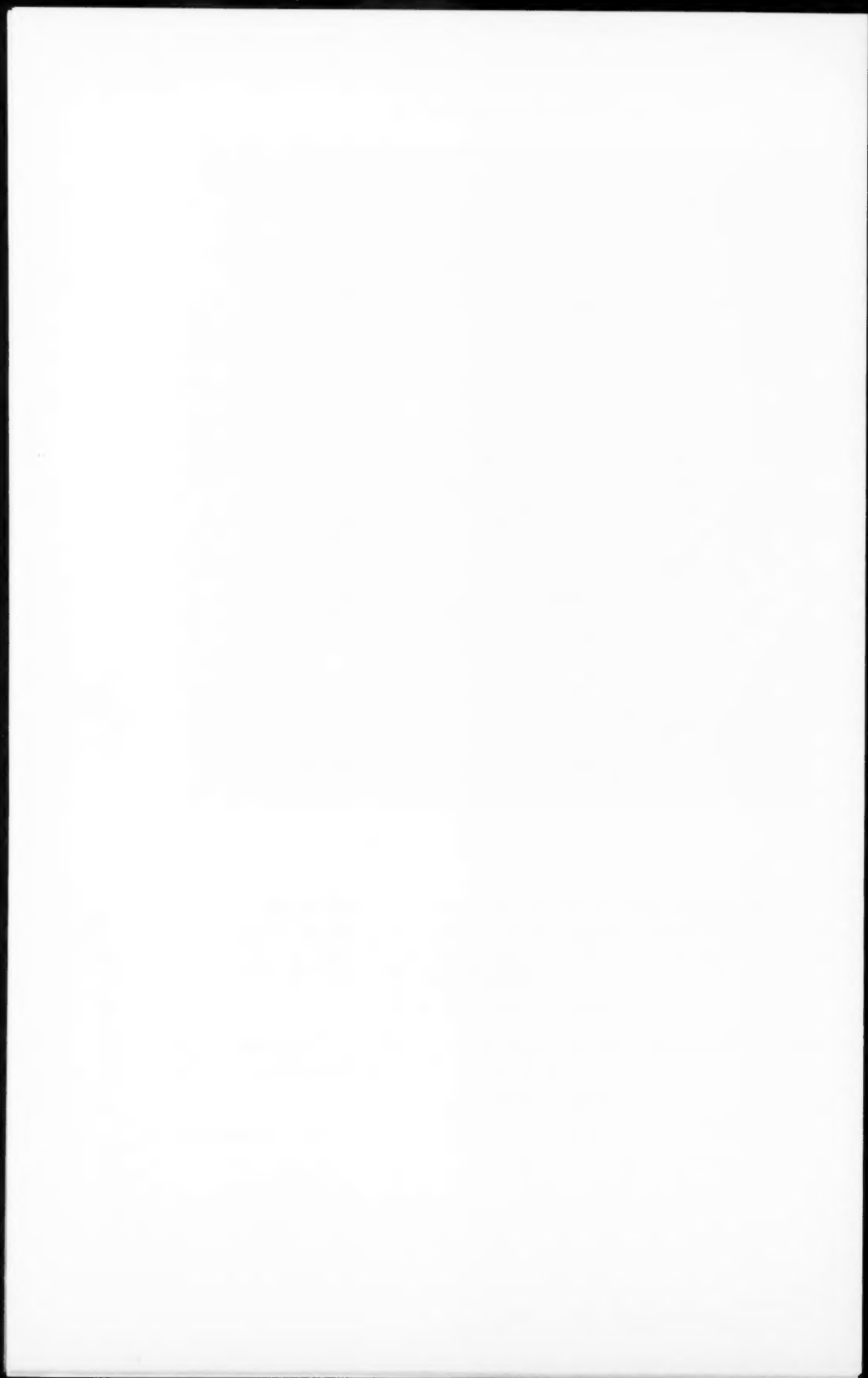
$$P^2 = 1, \quad PQP = QPQ$$

(implying  $Q^2 = 1$ ).

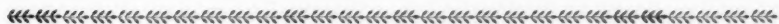
It may well be said that some kind of symmetry is the essential ingredient of all branches of mathematics. After a brief pause for questions and relaxation, Professors Opechowski and Wright will go still farther and tell us about its role in physics and chemistry. I will close with another quotation from the little book on *Symmetry* by Weyl (10, p. 5): "Symmetry, as wide or as narrow as you may define its meaning, is one idea by which man through the ages has tried to comprehend and create order, beauty, and perfection."

## REFERENCES

1. W. A. Bentley, *Snow crystals* (New York, 1931).
2. H. S. M. Coxeter, *Regular polytopes* (London, 1948).
3. H. S. M. Coxeter, M. S. Longuet-Higgins and J. C. P. Miller, *Uniform polyhedra*, Philos. Trans. Roy. Soc. London, A, 246 (1954), 401-450.
4. H. S. M. Coxeter and W. O. J. Moser, *Generators and relations for discrete groups*, Ergeb. Math., 14 (Berlin, 1957).
5. Sir Arthur Eddington, *The mathematical theory of relativity* (Cambridge, 1924).
6. R. Fricke and F. Klein, *Vorlesungen über die Theorie der automorphen Funktionen, I* (Leipzig, 1897).
7. Owen Jones, *Grammar of ornament* (London, 1868).
8. F. Klein, *Vorlesungen über das Ikosaeder...* (Leipzig, 1884; translated by G. G. Morris as *Lectures on the icosahedron*, London, 1913).
9. A. Speiser, *Theorie der Gruppen von endlicher Ordnung* (Berlin, 1924).
10. H. Weyl, *Symmetry* (Princeton, 1952).







## A SYMPOSIUM ON SYMMETRY

## II. Symmetry and Interaction between Elementary Particles

W. OPECHOWSKI

Presented by H. S. M. COXETER, F.R.S.C.

Towards the end of the last century, Maxwell's theory of electromagnetic phenomena, and in particular, that part of it which pertained to light and its ether endowed with all kinds of contradictory properties, appeared to many physicists to be very abstract, strange and almost incomprehensible. In view of Hertz's discovery of electromagnetic waves, they had to concede that Maxwell's theory was correct in its essentials, but to admit to understanding the theory was somehow too much for them. Lord Kelvin's many frantic attempts to "get" (as he said) the electromagnetic theory of light by making mechanical models are well known. Some other physicists went still further in this direction. In 1905, Pierre Duhem, a French theoretical physicist and historian of science, wrote about Sir Oliver Lodge's book on the subject (2): "Here is a book intended to expound the modern theories of electricity and to expound a new theory. In it there are nothing but strings which move around pulleys, roll around drums, go through pearl beads, carry weights; and tubes which pump water while others swell and contract; toothed wheels which are geared to one another and engage hooks. We thought we were entering the tranquil and neatly ordered abode of reason, but we find ourselves in a factory."

The attitude which Duhem derided seems to a modern theoretician not only old-fashioned, but quite impossible. And the reason for this extreme change in thinking is not the trivial fact that we know much more about the physical world than Lodge did, and that consequently an explanation of the electromagnetic phenomena in terms of engineering mechanics seems incongruous to us. The reason lies much deeper. I believe that nowadays many theoretical physicists would agree (although Kelvin and Lodge would not) that understanding the nature of the physical world can hardly mean anything else than (1) becoming thoroughly familiar with the mathematical structure of all those physical theories whose consequences we accept as being in agreement with experiment; and (2) seeing clearly which consequences depend on which features of that structure.

Around 1900, while endeavouring to achieve familiarity with the mathe-

mathematical structure of Maxwell's theory, Lorentz and Poincaré discovered a very strange symmetry of the basic equations of the theory. This symmetry we now call the "invariance of Maxwell's equations under Lorentz transformations," or "the invariance of Maxwell's equations with respect to the Lorentz group," or simply "Lorentz invariance."

Apparently neither Lorentz nor Poincaré was as convinced as many of us are today that every essential fundamental property of the mathematical structure of a successful physical theory (as Maxwell's theory was) expresses some fundamental property of nature. In fact, neither Lorentz nor Poincaré formulated the physical consequences of the invariance (symmetry) they had discovered, although Poincaré saw some of them very clearly. As is generally known, it was Einstein who, in 1905, was the first to do that. At the same time, he showed that if we are to avoid inconsistencies when interpreting experiments which involve both the electromagnetic and mechanical quantities, we have to modify the basic equations of mechanics in such a way as to make them also invariant with respect to the Lorentz group. He then performed the necessary modification, setting up what we call relativistic mechanics; the successes of this theory are known to everyone.

From what Lorentz and Poincaré did not do, and Einstein did do, we have learnt a lesson, and perhaps we have learnt it too well, as you will see later. In the first place, we dare not regard any fundamental physical theory as acceptable if it does not possess that basic symmetry, the Lorentz invariance. This applies, in particular, to all theoretical attempts to understand the behaviour of the numerous particles newly discovered. In the second place, we are very sensitive, so to speak, to any other symmetry our theories may have. We even look for symmetries where perhaps there are none.

It is now high time I said more explicitly what is meant by the statement that a theory shows some specific symmetry.

From the physical point of view, the essential characteristic of an object which is symmetric with respect to some operation is that there exists no method of ascertaining that the operation has actually been carried out on the object unless one has watched the operation being carried out.

If you leave a cube on your desk and somebody, during your absence, turns the cube  $90^\circ$  about an appropriate axis, you will not be able, under ideal conditions, to find out that he did so. Or if you leave several identical balls on the desk and somebody, during your absence, interchanges them, you will not be able to find that out either. This is so because, in the first example, the cube is symmetric with respect to twenty-four rotations of which one was carried out; and because, in the second example, identical objects are symmetric with respect to all permutations.

If the object is a physical theory and if we assert that the theory is symmetric with respect to some operations, or, as we would usually say, invariant with respect to some transformation of quantities occurring in

the theory, this assertion has the same meaning as in the two examples just mentioned. Mathematically speaking, the assertion means that the basic equations of the theory look exactly the same before and after the transformation.

Consider, for example, a theory which is invariant with respect to rotations in an ordinary three-dimensional space; and suppose that in that theory two mutually perpendicular vectors  $A$  and  $B$  play some part. Then the equation

$$A_x B_x + A_y B_y + A_z B_z = 0$$

will look exactly the same no matter how you choose the rectangular system of coordinate axes relative to which  $A_x, A_y, A_z$  are the components of  $A$ , and  $B_x, B_y, B_z$  are the components of  $B$ , provided you choose one among those coordinate systems which have a common origin and can be obtained from one another by mere rotation around the origin.

Rotations leave unchanged the squared distance from the origin,

$$x^2 + y^2 + z^2,$$

although they are, of course, not the most general linear transformations of  $x, y, z$  with this property.

Now, what Lorentz and Poincaré have noticed is that the basic equations of Maxwell's theory look exactly the same before and after any linear transformation of  $x, y, z$  and  $t$  ( $t$  being the time) which leaves unchanged the expression

$$x^2 + y^2 + z^2 - c^2 t^2,$$

where  $c$  was originally interpreted as meaning the velocity of propagation of light in the ether. This invariance of Maxwell's theory turns out to imply that the coordinate systems among which we may choose may not only be arbitrarily oriented in space as in a rotation invariant theory, but may also be in the state of translatory motion with respect to one another, provided the speed of the motion is constant. All this was quite clear to Poincaré. It was Einstein, however, who unambiguously showed that the concept of ether can be entirely eliminated from the theory, and that  $c$  must have the meaning of the velocity of light in empty space, this velocity being the same relative to all those coordinate systems.<sup>1</sup> This independence of the velocity of light of the coordinate system relative to which it is measured seems very strange at first sight, and it took some time before Einstein's conclusions became generally accepted.

In physical terms, the assertion that a theory is Lorentz invariant means then, as we have learnt from Poincaré and Einstein, essentially this: no experiment compatible with the theory could tell you that the space-rocket

<sup>1</sup>To what extent these revolutionary changes in physical theory are due to Lorentz and Poincaré and to what extent to Einstein is an historical question which is still a subject of controversy. See, for example, Sir Edmund Whittaker (19), and Max Born (1).

in which you are travelling in the intergalactic space has today a different constant velocity and a different constant orientation in space from the ones it had yesterday, unless you look at the stars through a window, and assuming of course that the change occurred when you were sleeping.

The qualification "unless you look at the stars" is, clearly, of paramount importance. It means that one must exclude all those experiments which involve physical systems to which the transformation has not been applied. But this is never, strictly speaking, possible. One may not look at the stars, but one cannot eliminate the very weak gravitational forces which are present even in the intergalactic space.

This is an aspect of the usual, fundamental difficulty of all physical theories to the extent they need the concept of a closed physical system. To put it differently, one can hardly attach a physical meaning to the assertion that a theory is Lorentz invariant, unless the theory is also invariant with respect to translations in space and translations along the time axis. In simple words, we have to assume that space has always been, is and will always be everywhere the same. This is a very strong assumption, but most theories of elementary particles make it. However, some of these theories are very satisfactory, so that we have at least an excuse to disregard gravitation, and the possible relevance of the structure of the universe, advocated by such people as Mach, Eddington and Milne. For similar reasons, nothing will be said in this talk about Einstein's General Theory of Relativity.

Consequently, whenever I say in this talk that a theory is Lorentz invariant I shall mean a theory invariant with respect to the transformations which leave

$$x^2 + y^2 + z^2 - c^2 t^2$$

unchanged, and, in addition to that, invariant with respect to translations in space and time.<sup>2</sup> Or, in other words, a theory invariant under transformations which leave

$$(dx)^2 + (dy)^2 + (dz)^2 - c^2(dt)^2$$

unchanged.

A point that I must emphasize because it will play a role later in my talk is that this definition includes the invariance with respect to what physicists usually call "space inversion" and "time reversal." Formally, time reversal means changing the sign of the time variable  $t$ . I will hardly mention time reversal in this talk, but certainly not because the subject is not interesting. "Space inversion" formally means the change of the sign of all three coordinates  $x$ ,  $y$ ,  $z$ . The space inversion when combined with an appropriate rotation is equivalent to a "mirror reflection," that is, to a transformation which changes the sign of one of the three coordinates

<sup>2</sup>However, the invariance with respect to time translations in the case of theories of particles that are "unstable" in the sense discussed below cannot obviously be exact.

leaving the other two unchanged. Hence an object symmetric with respect to space inversion and all rotations has a left-right symmetry. A similar statement holds for a theory invariant under these transformations. It follows that no experiment compatible with a Lorentz invariant theory could tell you that the space-rocket in which you are supposedly travelling underwent space inversion or mirror reflection. Of course, carrying out a mirror reflection on a space-rocket, or any other object, in the same sense in which one can turn the space-rocket is hardly feasible. This is why one usually explains the physical meaning of the assertion that a theory is invariant with respect to mirror reflection in a different, more realistic manner. According to this alternative point of view, the assertion has the following meaning: if a physical phenomenon is compatible with such a theory then, necessarily, the same phenomenon as observed in a mirror is also compatible with the theory, that is, it could actually occur, and not only as a mirror reflection of an actual phenomenon. In its literal sense, this physical criterion of the mirror invariance cannot obviously be applied to atomic phenomena. However, for macroscopic phenomena it makes perfect sense to speak about their reflection in a mirror. Now, all experiments on atomic phenomena reduce, in the final analysis, to observations of some simple macroscopic phenomena. Hence the criterion is more useful than it may seem at first sight. Of course, the invariance with respect to mirror reflection does not imply that the frequency of occurrence in nature of two phenomena which can be obtained from one another by mirror reflection is necessarily the same. This frequency may be determined by some other factors; in particular, by the structure of the universe as a whole.

All those Lorentz transformations which do not involve space inversion and time reversal are called the "proper" Lorentz transformations, and one speaks of the "proper Lorentz group." This is not a very precise statement, but it can easily be made precise. All Lorentz transformations as defined a moment ago form the "full Lorentz group."

As you may have noticed, the title of my talk is "Symmetry and Interactions between Elementary Particles." So far, I have been speaking about symmetry, and, in particular, the symmetry that every fundamental physical theory is supposed to have, Lorentz invariance. I shall now have to state a few things about elementary particles. Without discussing the very interesting question of what the adjective "elementary" means in this connection, I will simply enumerate the elementary particles, or rather their "families." After which, I will come to the question of the quantum mechanical description of their interactions, and to the "and" in the title.

Let us then have a quick look at the list of different families into which the particles known at present can be divided from the experimental point of view, which means, essentially, according to the order of magnitude of their masses (the electric charges seem to be always equal, in absolute

value, to the electronic charge, or to zero). For more details, see, for example (7).

The third column in the Table below gives the name of the family. A family may consist of several kinds of particles. For example, "electron" means a negative or a positive electron; "pions" may be positive, negative or neutral; "neutrinos" probably are of two kinds; "hyperons" consist of several kinds of particles; etc. The first column gives an idea of the masses of particles belonging to a family. In the fourth column the names given to groups of families are listed. The second column will be explained later.

The names of those particles which are stable are italicized. All remaining particles in the Table are unstable, that is, they have a finite "mean life-time," and many of them are just experimentally defined by their life-times and by the way they disintegrate into other (stable or unstable) particles. For example, a charged pion has a life-time of  $2.5 \times 10^{-8}$  second, and decays into a charged muon and a neutrino.

2180-2580	<i>F?</i>	Hyperons	} Baryons
1836	<i>F</i>	Nucleon ( <i>Proton</i> , Neutron)	
965	<i>B?</i>	Heavy Mesons	
264-273	<i>B</i>	Pion	} Light Mesons
207	<i>F</i>	Muon	
1	<i>F</i>	<i>Electron</i>	} Leptons
0	<i>F</i>	<i>Neutrino</i>	
0	<i>B</i>	<i>Photon</i>	

It is, I think, not surprising that these stable particles are the ones that have been known the longest. The neutron is almost stable; 13 minutes is a tremendously long time on an atomic time scale. As you know, atomic nuclei are just systems of protons and neutrons. Atoms, molecules and all that chemistry is about consist of nuclei and negative electrons, with photons as very frequent guests in this company.

On the other hand, hyperons and heavy mesons are newcomers. None of them has been known for more than ten years. They are often called "strange particles," and they are strange indeed. This is a point to which I will briefly come back later. Here, I will only mention that at least in one respect these particles are supposed not to be strange: in determining their masses, charges, life-times, etc., one assumes, of course, the validity of Lorentz invariant (that is, relativistic) mechanics and electrodynamics, and one uses a bit of quantum theory, just as much as one does for "ordinary" particles.

It was primarily to describe the interaction of electrons (free or bound in atoms, molecules, solids) with photons, that quantum mechanics was invented and developed. And, as a theory of electrons and photons, it is extremely reliable, in spite of some remaining difficulties.

However, it has always been clear that quantum mechanics is based on two more or less independent sets of assumptions. One set of assumptions



determines the nature of the general mathematical formalism of quantum mechanics and the physical interpretation of that formalism. (Let us call them Assumptions A.) The other set of assumptions has specifically to do with the interaction of electrons and photons and the way this should be introduced into the mathematical formalism (Assumptions B).

When we now turn to the problem of setting up a theory of the remaining particles in the list, then we obviously know in advance that we shall have to replace Assumptions B by something else, but we may still hope that Assumptions A are sufficiently general for our purposes, because they determine the general character of the mathematical formalism and not its details. The history of physics in the last twenty-five years shows that this hope seems to have been largely justified, and even in the case of strange particles it is not at all certain that the hope will have to be abandoned.

I want to discuss the part played by the invariance (symmetry) considerations in these two sets of assumptions on the example of the theory of beta decay. The simplest case of a beta-decay process is one in which a single free neutron disappears, and, in its place, a proton, a negative electron and a neutrino are created. The process thus involves two nucleons, and two leptons. I may mention in passing that this is the simplest case logically; from the experimental point of view it is a very difficult one, and it is only a few years ago that Snell and Miller (17), and Robson (15), a Fellow of this Society, succeeded in discovering the process, and they investigated it in detail. A process in which a proton, not free but bound in an atomic nucleus, disappears, and, in its place, a neutron, a positive electron and another kind of neutrino are created, also occurs.

I shall now briefly describe certain relevant features and consequences of Assumptions A, that is, those assumptions which determine the mathematical structure of quantum mechanics. I apologize in advance for a very fragmentary and distorted picture of the formalism of quantum mechanics, or the quantum theory of fields, as this most general form of quantum mechanics is often called.

To each kind of particle (not to each particle!) we assign in quantum mechanics a quantity usually called the "field" (or the "field operator"). One thus speaks of electron field, neutron field, photon field (which is essentially the electromagnetic field), etc. A field, which I will denote by the Greek letter  $\psi$ , may consist of several components, each component being a function of  $x, y, z, t$ . The several components of  $\psi$  satisfy, as functions of  $x, y, z, t$ , a set of differential equations which play a role analogous to that of equations of motion in classical mechanics. Of course, we require these differential equations to be Lorentz invariant. The equations thus remain unchanged under a Lorentz transformation, but the components of  $\psi$  in general, do change, just as the components of an ordinary vector change under ordinary rotations.

In mathematical language, one can say that the components of  $\psi$  transform linearly among themselves and in this way generate an irreducible

representation of the Lorentz group. That is, each field has well-defined "transformation properties" with respect to the Lorentz group. In other words, different kinds of particles can be characterized, although not uniquely, by their symmetry properties relative to Lorentz transformations.

In the case of electrons,  $\psi$  has four components which satisfy the so-called Dirac equation (I speak of the Dirac equation, in singular, because this is the established custom, but actually it is a system of four equations). The experimental evidence for the validity of Dirac's equation in this case is extremely convincing. With the other three kinds of particles which are involved in a beta-decay process, one had good reasons to assume, until a few months ago, that their  $\psi$ 's also satisfied the same Dirac equation as electrons do, and consequently had four components, except that the mass of a particle which occurs in the Dirac equation as a parameter is different in the four cases. However, this last circumstance does not affect the transformation properties of  $\psi$  which are the same in the four cases (although the case of zero mass is somewhat exceptional). Now, an important discovery of 1957 is that the neutrino field seems to be a different kind of  $\psi$  after all.

A field  $\psi$  assigned to certain kinds of particles may or may not be a measurable physical quantity. This is a question which I cannot discuss here. The essential point is, however, that all measurable quantities of a system of identical, non-interacting particles can be expressed in terms of  $\psi$  according to definite rules. Important measurable quantities are, for example, the number of particles the system contains, and the total energy of the system.

What I have just stated contains, strictly speaking, rather serious distortions of the truth. First, the  $\psi$ 's actually are not only functions of  $x, y, z, t$ , but also operators operating on certain abstract vectors in a Hilbert space. And what actually corresponds to measurable quantities are certain Hermitian operators defined in terms of the  $\psi$ 's. Such things as the number of particles or their total energy are only eigenvalues of those Hermitian operators. But I cannot, of course, go into such technicalities. However, I must mention a second point (which is an omission rather than a distortion). In a system of  $N$  identical particles, one would expect that the group of all permutations of  $N$  objects (the so-called "symmetric group") would somehow play an essential part in the corresponding quantum mechanical formalism. In fact, it does. The part played by the symmetric group manifests itself in the so-called "commutation relations" imposed on the  $\psi$ 's regarded as operators. The curious thing is, however (and here I must again use technical mathematical terms for a moment), that only the two simplest representations of the symmetric group seem to matter from the physical point of view: the so-called "symmetric" representation, and the "anti-symmetric" one, and correspondingly there are only two types of commutation relations in the quantum mechanical formalism. This means that, according to that formalism, all possible kinds



of particles can be divided into two classes as far as the symmetric group is concerned. Those kinds of particles which, so to speak, "belong" to the anti-symmetric representations are called "fermions" (from Fermi's name) and those which "belong" to the symmetric representation are called "bosons" (from Bose's name). Fermions and bosons each behave in quite a different and characteristic way in many phenomena (for example, properties of a fermion or boson gas are quite different). I have indicated by  $F$  or  $B$ , in the second column of the Table, which particles are fermions and which are bosons.

Turning now to the problem of interaction between particles, I must first say that according to the quantum mechanical formalism and its physical interpretation, particles of any kind are stable unless they are in interaction with some other kind of particles. But as soon as there is an interaction between different kinds of particles, such particles may be annihilated or created. The best-known example is, of course, absorption (annihilation) and emission (creation) of photons by electrons.

An interaction between particles is taken into account in the mathematical formalism of quantum mechanics by introducing a corresponding Lorentz invariant "interaction operator," which is an expression involving the different fields assigned to the several kinds of particles supposed to be in interaction. Once the interaction operator is known, all questions of experimental interest can, in principle, be answered according to well-defined mathematical procedures. You have noticed the qualification "in principle," and you know that this often means that some questions cannot be answered.

The choice of the interaction operator involves what I have called Assumptions B.

In the electron-photon interaction, the correct interaction operator could easily be guessed from our knowledge of macroscopic electromagnetic phenomena, the mathematical description of which is, so to speak, by definition Lorentz invariant. For example, we have long known that the force between two small, charged macroscopic bodies is inversely proportional to the square of their mutual distance; we have also long known what the intensity distribution of the electromagnetic radiation emitted by an antenna is.

For all other particles listed in the Table we have nothing comparable to go by. No similar macroscopic phenomena caused by other interactions are known or even believed to exist. (I disregard gravitation which seems to be quite irrelevant to the atomic and nuclear phenomena at the present degree of experimental accuracy.)

In particular, one has no such clues for that interaction which leads to the beta decay of the neutron (and of systems of neutrons and protons: I mean, atomic nuclei). As a result the beta-interaction operator on which the theory of beta decay is based has been almost entirely determined by the requirements of Lorentz invariance. I say "almost entirely" because,

historically, Fermi, who in 1934 put forward the theory, was partly guided by analogy with electron-photon interaction. However, the analogy is perhaps superficial.

This analogy makes us look for all those Lorentz invariant expressions which are (using again formal mathematical language) bilinear with respect to the  $\psi$ 's of the nucleons, and also bilinear in the  $\psi$ 's of the leptons.

Let us consider this more in detail. If one first introduces a less stringent requirement, namely, that the expressions be invariant with respect to the proper Lorentz group (that is, excluding the space inversion and time reversal) then there are ten such invariants (let us denote them by  $\Gamma_k$ ), and the most general interaction operator  $H$  for the beta or Fermi interaction (as this interaction is called) is the sum

$$H = \sum_{k=1}^{10} C_k \Gamma_k$$

where the  $C_k$  are just ten arbitrary complex numbers, which are, so to speak, symbols of our ignorance.

If one next demands that the interaction operator be also invariant with respect to time reversal, then—it turns out—the  $C_k$  become real numbers. This reduces our ignorance by a factor 2, because each complex number is a pair of real numbers.

Finally the requirement that the interaction operator be also invariant with respect to space inversion makes five of the ten real constants vanish, or, in other words, five of the ten  $\Gamma_k$ 's are not invariant with respect to space inversion. In this way one concludes that the invariance with respect to the full Lorentz group makes the theory of beta decay depend on five arbitrary parameters. And, of course, one demands as usual that the theory be invariant with respect to the full Lorentz group.

By comparing the predictions of the theory containing these five parameters with the corresponding experimental results one can determine their numerical values. And this is what one has been busy doing for the last twenty years, as the problem is a very complicated one. The theory has been getting gradually well established from the experimental point of view, and no one expected any world-shaking developments in this branch of physics.

I don't know how one defines a "world-shaking" event, but I am sure that when a discovery involving pure science with no prospect of application gets a headline on the front page of the *New York Times* it is world-shaking. On January 16 of this year, the discovery that the Fermi interaction operator is, after all, not invariant with respect to the space inversion was announced in the *New York Times* in much more spectacular language than I am using here. There is no semi-scientific or scientific magazine which has not devoted some space, since that time, to the now celebrated "non-conservation of parity" as the discovery is generally called. Many among you must have seen diagrams illustrating the gist of these history-making

experiments. Faithful to the rather abstract, non-pictorial way of presenting the symmetry properties of physical theories, which I have been following in this talk, I am not going to draw diagrams. I prefer, for a change, to make somewhat vague statements about what the theory precisely is rather than precise statements about diagrams that vaguely illustrate the theory.

First of all I want to emphasize that the discovery was not accidental. It was a result of a conscious, well-organized, tremendous effort of a group of American physicists, started by two theoreticians Lee and Yang, for whom this is not the first brilliant achievement. I cannot mention all the names of the many experimentalists involved, but I will make the exception for a lady, Miss Wu, who has been very well known for her work on beta decay, and who played a decisive part in one of the experiments.<sup>3</sup> In a way, one could thus say that this was a Chinese-born American discovery. It may be mentioned that Salam (16) in England and also an eminent Russian theoretician, Landau (8), have somewhat later, but independently, expressed similar ideas to those of Lee and Yang. Almost no relevant experimental work was set off by Landau's ideas in Russia, which prompted a Belgian friend of mine to write to me that, after all, the United States this time gives a good example of collective planning, and the Soviet Union of free enterprise. "Somme toute, les Etats-Unis nous donnent un bel exemple d'effort collectif, l'Union Soviétique de prestation individuelle."

Doubts concerning the universal validity of the invariance with respect to space inversion apparently arose in Lee and Yang's minds when they were trying to understand some of the peculiarities of the "strange particles." This led them to examine very carefully the existing experimental evidence for the validity of that invariance in all well-established theories describing the interactions of the better-known particles. Their conclusion, published last October (9), was that the evidence was good except in the theory of beta decay. They found that, somewhat unexpectedly, all the numerous experiments related to beta decay tell us absolutely nothing about the invariance of the theory with respect to space inversion.

The mathematical reason for that can be stated very simply. Let us suppose that the theory is not invariant. Then, as we have seen before, the interaction operator is a sum of five terms which are invariant and another five terms which are not. To emphasize this let us rewrite the operator as follows:

$$H = \sum_{k=1}^5 C_k \Gamma_k + \sum_{k=1}^5 C'_k \Gamma'_k$$

where the first sum contains the five invariant expressions, and the second sum contains the five that are not invariant. Yang and Lee have shown

<sup>3</sup>On second thought, I feel I should mention that in addition to the experiment of Miss Wu and her collaborators (21) three other experiments on the non-conservation of parity have been carried out almost simultaneously; see (3), (4) and (14).

that the results of all experiments performed until that time depend only on expressions of the form

$$C_k^2 + C_k'^2 \quad \text{or} \quad C_k C_l + C_k' C_l'$$

so that they do not make it possible for us to determine primed and unprimed constants separately. In other words, we cannot conclude if the second, non-invariant sum is zero or not.

Next, Lee and Yang have indicated some experimental effects in beta decay which solely depend on mixed products of primed and unprimed constants, for example,  $C_k C_k'$ . The mere existence of such effects would mean that at least some of the primed constants are different from zero, which in turn would mean that the theory is not invariant with respect to space inversion. And, as we have seen, the big discovery is that these effects do exist.

Finally, as soon as the existence of these effects became fairly certain, Lee and Yang (10) put forward a modified theory of beta decay in which the field  $\psi$  assigned to the neutrino has only two components, and satisfies an equation which is, of course, no longer invariant with respect to space inversion. The equation (due to Weyl) has been known for many years, but no one took it seriously, so strongly were we addicted to the universally valid symmetry principles.

In terms of the interaction operator  $H$  the new theory means that the corresponding primed and unprimed constants differ at most in sign, that is,

$$C_k = \pm C_k'$$

You may recall that there are probably two kinds of neutrinos: the neutrino which accompanies the emission of a negative electron in beta decay, and the neutrino which accompanies the emission of a positive electron. Now, neutrinos as all other fermions have necessarily an intrinsic angular momentum, the "spin." According to Lee and Yang's new theory, the direction of translatory motion of a neutrino determines uniquely the sense of its spinning. In other words, the two kinds of neutrinos are related like a right-handed and left-handed screw. If Lee and Yang's theory is correct (which it very well may not be!), the interpretation of all phenomena in which neutrinos take part will have to be modified.

In this final part of my talk, I want to speak about the so-called "conservation laws" and their close connection with the invariances of a theory. We say that a physical quantity defined for a closed physical system is "conserved" when its value does not change in course of time. It has been known long before the advent of quantum mechanics that such a close connection exists. For example, in classical Newtonian mechanics the invariance under translations in space and time implies the conservation of momentum and energy. Similarly, invariance under rotations leads to the conservation of the angular momentum. Like statements hold true in classical Lorentz invariant mechanics.

In quantum mechanics we have again a comparable situation. However, whereas in classical mechanics the invariance under space inversion and time reversal does not lead to anything new, in quantum mechanics it does. In particular, the invariance of a quantum mechanical theory under space inversion implies the existence of a physical quantity (represented, as usual, by an operator) which is conserved. This quantity is called "parity." This is why the discovery that the beta interaction is not invariant under space inversion has often been referred to as the discovery of the non-conservation of parity.

The importance of conservation laws in predicting and interpreting the results of experiments is well known. The conservation laws are especially helpful in processes involving those elementary particles whose properties are still obscure. In fact, the discovery of a new particle usually means that an experiment involving well-known particles cannot be interpreted without violating some conservation laws, unless we assume that a new, unknown particle is also involved.

In quantum mechanics the general consequences of conservation laws for simple processes are formulated as the so-called "selection rules." Thus one speaks of the angular momentum selection rules, parity selection rules, etc. Because of the connection between invariance and conservation laws, each selection rule expresses a mathematical property of the group of transformations with respect to which the theory in question is invariant. For the mathematicians, I may perhaps mention that a selection rule is just a theorem concerning the way a Kronecker product of two irreducible representations of the group reduces into irreducible representations.

We have seen that all theories of elementary particles are invariant with respect to the proper Lorentz transformations, and all of them are also invariant with respect to space and time reflections, except the theory of beta decay, probably the theories of all phenomena which involve neutrinos, and perhaps the theories of some among the strange particles. We have also seen that all those theories are invariant under all permutations of identical particles.

The question arises whether there are any other transformations under which these theories are invariant. In fact, such transformations exist. This is very important because they lead to additional conservation laws, and hence to additional selection rules. However, these transformations are not transformations of  $x$ ,  $y$ ,  $z$ ,  $t$ . They are defined as certain transformations of the field operators  $\psi$ . We have seen that the Lorentz transformations which are transformations of  $x$ ,  $y$ ,  $z$ ,  $t$  generate transformations of the components of  $\psi$ . The new transformations I am now speaking about are defined directly in terms of the  $\psi$ 's.

I will mention one such transformation which has been called by Kramers "charge conjugation"; it is also called "particle-anti-particle conjugation." It is a transformation which for charged particles corresponds to changing the sign of all charges, but can also be defined for neutral particles. "Charge

conjugation" transforms a negative electron into a positive electron and vice versa, etc., and, in general, a particle into its "anti-particle." The consequences of the invariance with respect to "particle-anti-particle" conjugation have been thoroughly described by science-fiction writers, and I do not intend to compete with them. There is, however, one point worth emphasizing: although we have succeeded in producing all kinds of anti-particles, for example, positron, anti-proton, anti-neutron, etc., we have not yet succeeded in producing one single anti-atom let alone a piece of an anti-solid.

Just as the invariance with respect to space inversion means that there exists an operator  $P$  which is "conserved," so the invariance with respect to charge conjugation means that an operator of charge conjugation  $C$  exists which is conserved. This leads to selection rules which are useful for disentangling experimental data involving production and decay of particles.

It is fairly certain that in beta decay not only is there no conservation of parity, but also there is no conservation of charge conjugation.

However, the theory of beta decay may very well be invariant under a transformation which consists in carrying out space inversion and charge conjugation one after another. Whether it is actually so is one of the several important questions which will have to be answered in the near future. The validity of that combined invariance would, for example, mean that what is left-hand side for us would be right-hand side for "anti-human beings" in their anti-world, even when questions concerning beta decay are considered.

There is a peculiar interdependence of the three transformations, charge conjugation, space inversion, and time reversal, which finds its expression in a following theorem of a very general validity: if a theory is invariant with respect to the proper Lorentz group, then it is necessarily invariant with respect to a transformation which consists in carrying out the charge conjugation, space inversion, and time reversal one after another, in any order. This theorem was first known as the Lüders theorem (11), then the Pauli-Lüders theorem, then the Schwinger-Pauli-Lüders theorem, and as the importance of the theorem increases so does the number of names attached to it. Some people call it simply the CPT-theorem. When one takes into account the non-conservation of parity, the theorem implies that the correct theory of beta decay cannot be invariant under both charge conjugation and time reversal.

We have seen that the invariance of a theory with respect to some transformations always means that there correspond to those transformations quantities which are conserved. Is the converse assertion also true? If we know about a physical quantity that it always has, for any closed system, a value independent of time, can we conclude that the satisfactory theory for the system must necessarily have some hidden invariance which corresponds to the conservation of the quantity in question? Or, more



succinctly, is a Conservation Law always equivalent to a Symmetry Principle? The answer to this question is not known.

Consider, for example, the electric charge which seems to be rigorously conserved in all physical processes. One can define a transformation, the so-called "gauge transformation," and the invariance of the theory with respect to this transformation has certainly to do with the conservation of charge. But there are still some obscure elements in this relation.

There is another conservation law, formulated a few years ago by Wigner (20), which in its most general form asserts that the number of baryons minus the number of anti-baryons remains constant.<sup>4</sup> It is fairly well established in the special case of nucleons. No one has yet observed a creation of a nucleon without simultaneous disappearance of another nucleon. Positive and negative protons always disappear together, etc. We do not know the symmetry principle, if any, which corresponds to this law of conservation of baryons. But there have been attempts to find such a principle.

Finally I would like to mention still another conservation law which has been formulated in the course of numerous attempts to understand why certain processes involving strange particles do occur, and others do not or at least very rarely. The fact that a process does not occur, although it is logically possible, usually means in physics that its occurrence would violate some conservation law. In other words, the process in question is forbidden by a selection rule. Now, Gell-Mann (5) and, independently, Nakano and Nishijima (12) in 1953 succeeded in assigning to each kind of strange particle a quantity (a very simple quantity, just a dimensionless integer) the conservation of which seems to explain quite satisfactorily the occurrence or non-occurrence of processes involving strange particles. The quantity has been called "strangeness," and one speaks of the "conservation of strangeness." The conservation of strangeness is not supposed to be a universally valid conservation law. Just like the conservation of parity it is violated in some processes.

Does the conservation of strangeness correspond to some symmetry principle? Supposing that the fields  $\psi$  of the strange particles depend on certain new variables in addition to  $x, y, z, t$ , it is, in fact, possible to construct mathematical formalisms corresponding to the conservation of strangeness. The idea of introducing additional variables is not new in quantum mechanics. However, it has never been taken as seriously as it is now. This new trend has its origin in the work of Pais (13; 6) in 1952 and 1953 (the investigations of Gell-Mann, and of Nakano and Nishijima were also strongly influenced by Pais's ideas). Many theoreticians invent new abstract spaces (introducing new variables means just that), consider all sorts of transformations in these spaces, and try to guess how all that could

<sup>4</sup>Having mentioned Wigner's name, I must add that no one else has contributed as much to the understanding of the part played by the theory of groups (that is, by the symmetry considerations) in quantum mechanical theories as he has.

possibly be interpreted in terms of the scanty experimental information about the strange particles.

I suppose some physicists would say about these attempts something similar to what Lord Kelvin (18) said seventy years ago about Maxwell's electromagnetic theory of light: "I want to understand light as well as I can without introducing things that we understand even less." I must confess that Kelvin's remark has always puzzled me. How can we ever understand anything really new, different, strange, without first introducing things that we hardly understand at all?

## REFERENCES

(Only some recent papers are listed below)

1. Max Born, *Physics in My Generation* (London and New York, 1956).
2. Pierre Duhem, *The Aim and Structure of Physical Theory* (Princeton, N.J., 1954), p. 70.
3. J. I. Friedman and V. L. Telegdi, *Phys. Rev.*, **105** (1957), 1681.
4. R. L. Garwin, L. M. Lederman and M. Weinrich, *Phys. Rev.*, **105** (1957), 1415.
5. M. Gell-Mann, *Phys. Rev.*, **92** (1953), 833.
6. M. Gell-Mann and A. Pais, *Proc. Glasgow Conf.* (1954), 342.
7. M. Gell-Mann and E. P. Rosenbaum, *Scientific American*, **197** (1957), 72.
8. R. Landau, *Nuclear Physics*, **3** (1957), 127.
9. T. D. Lee and C. N. Yang, *Phys. Rev.*, **104** (1956), 245.
10. T. D. Lee and C. N. Yang, *Phys. Rev.*, **105** (1957), 1671.
11. G. Lüders, *Det Kong. Danske Videnskabernes Selskab, Mat.-fysiske Meddelelser*, **28** (1954), no. 5.
12. T. Nakano and K. Nishijima, *Prog. Theor. Phys.*, **10** (1953), 581.
13. A. Pais, *Physica*, **19** (1953), 869.
14. H. Postma, W. J. Huiskamp, A. R. Miedema, M. J. Steenland, H. A. Tolhoek and C. J. Gorter, *Physica*, **23** (1957), 259.
15. J. M. Robson, *Phys. Rev.*, **78** (1950), 311.
16. A. Salam, *Nuovo Cimento*, **5** (1957), 299.
17. A. H. Snell and L. C. Miller, *Phys. Rev.*, **74** (1948), 1217.
18. W. Thomson, *Lectures on Molecular Dynamics and the Wave Theory of Light* (Baltimore, 1884), p. 270.
19. Sir Edmund Whittaker, *History of the Theories of Aether and Electricity*, II (New York, 1954).
20. E. P. Wigner, *Proc. Am. Phil. Soc.*, **93** (1949), 521.
21. C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, *Phys. Rev.*, **105** (1957), 1413.



## A SYMPOSIUM ON SYMMETRY

## III. The Symmetry Sense in Chemistry

GEORGE F WRIGHT, F.R.S.C.

This paper is not written to tell you about chemistry but, rather, to tell you about chemists; especially about their intuitive use of symmetry concepts. I hope that you will not feel you must comprehend the chemical detail, especially since much of it is unfamiliar even to chemists, if they do not happen to be interested in the organic chemical field. As I understand it, this symposium was organized that we might acquaint one another with the use to which the symmetry sense is applied in the various basic scientific disciplines. Perhaps it would be most profitable if you would regard this essay as an evaluation of myself, an average chemist, rather than as an exposition.

Some may be surprised when I define chaos as a highly symmetrical state. In making this definition I am thinking as a crystallographer, about substances which pass abruptly from the centrosymmetric triclinic form through monoclinic, then rhombic and tetragonal, finally to the isometric cubic form. But it is not these symmetry variations that are of principal interest, but rather the fact that in this sequence towards the highest symmetry, the cubic, the crystalline repeating unit becomes larger and larger. I can then melt the substance and attain the symmetry of the beaker which contains it. Then I can vaporize it and begin to approach the symmetry of chaos.

Ordinarily we don't think of symmetry in this sense because it isn't useful to us. In our region of the universe the symmetry we enjoy is that which gives us energy in consequence of a particulate orderliness and we say

$$F = H - TS$$

that is, available energy,  $F$ , is a consequence of the particulate order factor, or enthalpy,  $H$ , minus the disorder factor involving particulate interaction  $S$  multiplied by particulate motion  $T$ . It is implicit in this statement that the system we are considering is in equilibrium, or as an approximation, in a steady state.

In that part of chemistry where equilibrium reigns it is useful to define matter in terms of its symmetry, although the concept leads to error if these ideas of symmetry are based on three-dimensional space. But in simple

instances the concept is workable. For example, we know that the element carbon has a double electronic charge ( $2S^2$ ) almost spherically disposed about the nucleus, and two more electron charges disposed antispherically ( $2P^2$ ) about the nucleus. Since the  $S^2$  charge is helium-like in its symmetry, one would expect that this charge would be unaltered in the presence of hydrogen atoms. Instead, the two  $P$  orbital charges would couple with the 1  $S$  hydrogen charge so that an approximation to more helium-like symmetry would obtain by combination to the molecule, methylene.

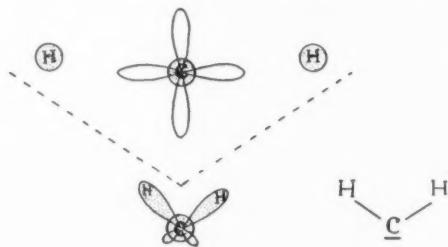


FIGURE 1

The indistinct picture of methylene (Fig. 1) has been idealized at the right to indicate the  $SP$  orbital of  $C-H$  as a "bond" connecting these atoms and a bar to indicate the two  $S$  electrons which are paired because of "opposite spin." Indeed this molecule, methylene, exists at elevated temperatures, but when chaotic symmetry decreases with decrease in temperature, a greater symmetry is achieved if the  $S$  and  $P$  orbitals of the carbon atom degenerate to four  $Sp^3$  orbitals of equal realm. These orbitals are best pictured in terms of their interaction with four hydrogen atoms to

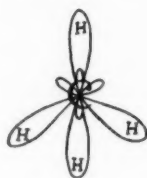
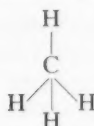


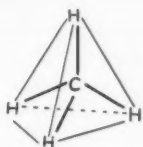
FIGURE 2

give methane (Fig. 2). This degenerate orbital system about the carbon atom now may couple with *four* hydrogen to give a new symmetry element of eight electron charge, the tetrahedral methane with hydrogen bonded fourfold to carbon.

The term "bonded" used in the last sentence is a questionable one. I shall accentuate the error by trying to improve my messy picture by the familiar one of apparent symmetry.



This one is much less messy, but it is also much less truthful. For example, the hydrogen atoms from which this methane was created had another choice; they could have achieved helium-like symmetry as hydrogen molecules. But this tendency could not entirely have ceased; so we should add additional "bonding."



But the hydrogen molecule tendency should lead to interaction of hydrogen molecules and these to the original methylene tendency, and so on. We have thus come back to the messy picture with the realization that the real symmetry is not conceivable in terms of the three-dimensional physical world about us. Physicists and chemists recognize the philosophy in the idea that methane is not a composite of bi-atom orbitals, but instead has a molecular orbital, or charge realm, to which I sometimes refer as the "symmetrical soul of methane."

Unfortunately it is difficult to ascribe dimensions to souls or, at least, to molecular orbitals. Consequently it has become popular to approximate the molecular orbital as a summation of atomic orbitals plus all of their interactions that are not too complicated or difficult to calculate. Of course, this is the scientific method. It is approximation by comprehension of the parts (derived from isolated ideas or isolated experimental observations) when one cannot comprehend the whole. The scientist dares to deal with partial truth in the expectation that impact with fact or with other partial truth at this vulnerable low level of reason will prevent a flight into absurdity. I have no quarrel with this methodology which is at least as old as Socrates. Indeed, I am inclined to laugh at those self-ordained priests of learning, the scholars, who cannot seem to comprehend this deliberate limitation of reason; who seem to derive their concepts of science from novels like *Frankenstein*. But do I laugh at myself when I use the naïve scientific approximations? I must admit that scientists often forget to be humble about their limitations as human beings. I will affirm that most error in philosophy, natural or human, stems from an insufficient sense of humour. Of course, I am merely paraphrasing the words of men like Kierkegaard, and even Nietzsche in his *Gay Wisdom*. Every vision of intelligence should be inspected closely in the mirrors of intelligence.

With a sense of humour we may safely use that imaginary figment of atomic interaction called the chemical bond. We may even elaborate upon it by using "hybridized" or "degenerate" or, worst, "resonating" atom orbitals. Actually, tongue in cheek, we will doubt that a soul can be compounded from a series of sub-souls. We are not so silly that we believe methane to be the symmetrical combination of one carbon and four hydrogen atoms which is suggested by our simple intuitive ideas of symmetry. Nevertheless, with healthy caution we can discover secrets of nature by application of intuitive symmetry concepts such as the simple or degenerate chemical bond. Organic chemists are especially prolific in such discoveries. As an example I shall describe Mr. H. Sawatzky's work at Toronto, not to convey the details of the chemistry so much as to show how the organic chemist uses his symmetry sense.

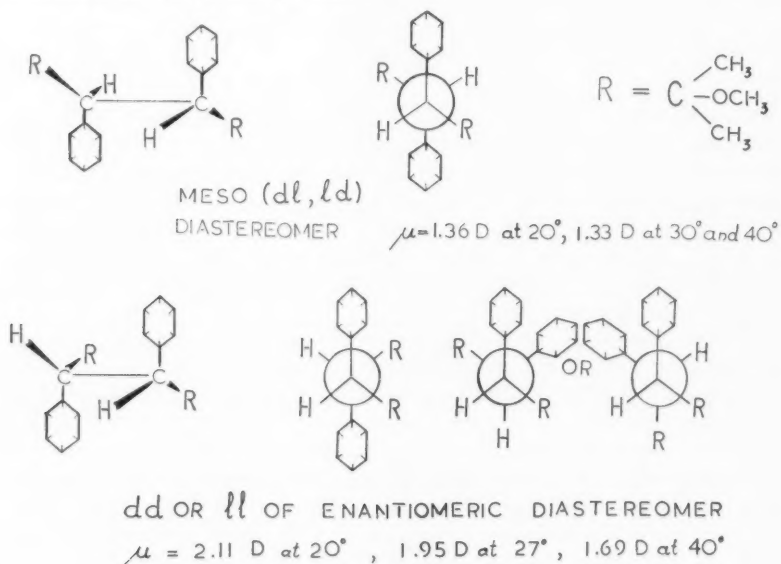


FIGURE 3

The structure shown in Figure 3 represents ethane,  $C_2H_6$ , in which one of the hydrogen atoms on each carbon has been replaced by a phenyl<sup>1</sup> group and another of the hydrogen atoms of the ethane has been replaced by a methoxypropyl group which is called "R". The actual composition of these substituent groups is unimportant; only their bulk and their force fields are significant to the argument. Observe that all four attach-

<sup>1</sup>The symmetrical phenyl group is pictured as a hexagon in which the fourth orbital of each tetravalent carbon atom is directed toward the centre of the hexagon. Thus six short radial lines so directed indicate the partial molecular orbital called a "pi sextet."

ments to each carbon of the ethane are different. Now we shall disregard influences outside the molecule and shall even consider separate regions of the molecule (which strictly we have no right to do) as centres of asymmetry. Then the relative arrangement of groups may be defined: looking in one direction down the pivot bond of the ethane one sees "phenyl, R and H" counter-clockwise (left-handed or *levo*), while in the other direction one sees "phenyl R and H" as clockwise (right-handed or *dextro*). We can call the combined relationship "*ld*," and since the groups on each carbon are the same we can turn the molecule through  $180^\circ$  and call it *dl*; actually we call the whole system a meso diastereomer. By contrast, if the other diastereomer shown in Figure 3 is viewed both ways from the pivot bond, one sees "phenyl R and H" both counterclockwise; so we call it the *dd*-diastereomer, realizing that an *ll*-diastereomer is equally probable on account of the symmetry of synthesis. That is to say, the system out of which these diastereomers are evolved has no unique sense of direction. Combinations of *d* and *l* parts may give rise to an equal distribution (*ld*, *dl*, or meso) within one diastereomeric molecule or else to an equal distribution (*dd* or *ll*) within those molecules which comprise the other diastereomer. Of course the two diastereomers will not be present in equal amounts because their energies are different.

These two diastereomers, of configurational difference, are quite different in physical and chemical properties. The structures of the two would be defined completely if there were free unobstructed rotation about the pivot bond. However, this assumption of free rotation has been tested in other substances such as 1,2-dichloroethane and found to be false. In order to demonstrate the effect of hindered rotation it is instructive to view the entire molecule end-on, to discover space-chemistry differences which Professor Melvin Newman calls rotamers. The end-on picture of the meso diastereomer (where the spider represents the first observable carbon atom and its substituents and the circle represents the rearward carbon atom with its substituents) is written with all the groups staggered, and opposed. Other rotamers could be drawn; but let us predict that only this rotamer will exist because of its symmetry.

By contrast the first rotamer of the *dd, ll*-diastereomer has only an oblique plane of symmetry. Rotation through  $120^\circ$  does not appreciably improve this symmetry, nor does another  $120^\circ$  rotation. Of course, an infinite number of rotamers intermediate between these staggered forms might be postulated, but bulk interference would render them improbable when the molecules were in their lowest energy states. It is evident that this *dd, ll*-diastereomer is uneasy with respect to the symmetry of these lowest energy states. Indeed, it can become more symmetrical only by overcoming the bulk hindrance to relatively free rotation.

Now let us couple each of these diastereomers to an electrostatic field. That is to say, let us determine the dielectric constants of the free substances in order to find which one will align itself most strongly with the field to

minimize the energy of the coupled system. Intuitively we know that the more symmetrical (meso) diastereomer orients less easily, and in fact we find that its dipole moment, though not zero because of moments outside the pivot linkage between the central carbon atoms, actually is less than that of the *dd,ll*-diastereomer. Moreover we find that the observed moment is experimentally invariant with respect to temperature, thus justifying the intuitive reluctance to depict more than one rotamer. By contrast the *dd,ll*-diastereomer shows a marked decrease in dipole moment with increased agitation at higher temperatures. This is a symptom of the unsymmetrical uneasiness predicted upon examination of the rotameric forms. Thermal agitation has removed the inhibition to rotation which the organic chemist calls steric interference by the bulky groups. Thus some intuitive ideas of symmetry have made it possible to predict which of the diastereomer is meso and which is *dd,ll*.

Lest we become too smug in this success, let us examine an instance in which the intuitive idea of symmetry fails. According to the orthodoxy of atomic structure divalent mercury possesses a pair of *S* orbital electrons in its valence shell and therefore should form linear C-Hg-C ( $\mu = \text{zero}$ ) in an organo-metallic compound like diphenylmercury (Fig. 4). To judge from

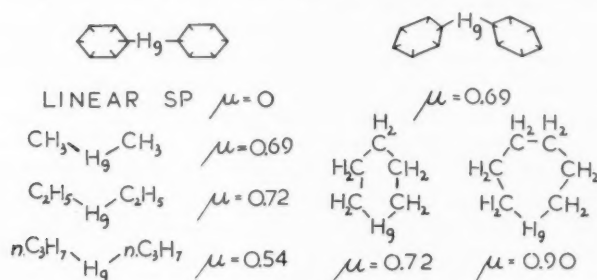
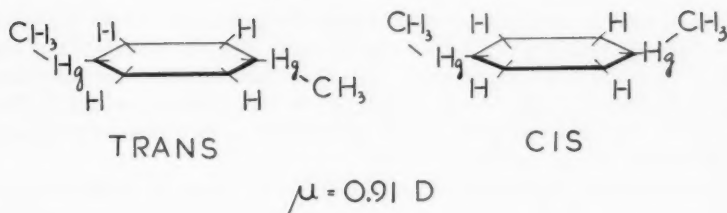


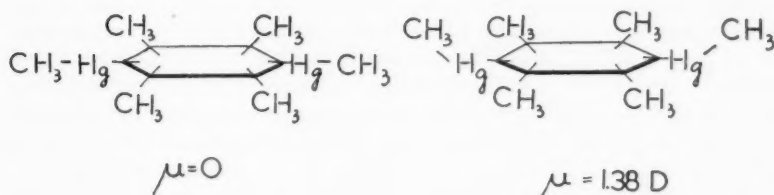
FIGURE 4

literature of chemistry it was a heretical affront when Hampson in 1934 found the dipole moment of diphenylmercury to be 0.69 D, thus indicating a bent molecule. Various attempts were made to do away with the heresy. These attempts finally culminated in an elaborate theory which explained away the experimental polarization as field distortion of constituent atoms rather than by orientation of diphenylmercury in the field. The alarm then subsided until we devised at Toronto a method of determining atomic polarization, and we found that diphenylmercury had none.

The workers in this field have not evinced pleasure at our renewal of the apparent heresy, but Sawatzky, who seems to be able to take his intuition or leave it, has strongly strengthened the case for the bent molecule. First he has shown (Fig. 4) that dimethyl-, diethyl- and dipropyl-mercury all have moments of the same order of magnitude as the moments of mercura-



## 1,4 DIMETHOMERCURIBENZENE



## BIS - METHOMERCURIDURENE

FIGURE 5

cyclohexane and mercuracycloheptane where, because of the cyclic structure of the C-Hg-C, linkage cannot reasonably be linear.

In a more elegant proof (based on the intuitive idea of symmetry) Sawatzky has synthesized dimethomercuribenzene and dimethomercuridurene, shown in Figure 5. If the C-Hg-C linkage were linear these substances should have zero moment. The experimental value of 0.91 D for dimethomercuribenzene shows that the two methomercuri groups are not linearly opposed. A separate determination of the group moment of the methomercuri group (0.74 D from the moments of bromo-benzene and 1-bromo-4-methomercuribenzene) shows that the deviation from C-Hg-C linearity is  $56^\circ$  if one assumes a simple harmonic distribution (free rotation) between the extreme opposition of the methomercuri groups (trans position, zero moment) and the extreme supporting contributions of these groups in the so-called *cis*-conformation. The angle  $\theta$  is obtained from the relationship

$$\mu_{\text{resultant}} = \mu_{\text{group}} \cdot \sqrt{2} \cdot \sin \theta.$$

Now it is of interest to compare this result with the one obtained for dimethomercuridurene, which contains four methyl groups known to hinder rotation in the 1 and 4 positions. The experimental moment (1.38 D)



can be related to the group moment and C-Hg-C angle used above by discarding the factor  $\sqrt{2}$ . This means, not only that the C-Hg-C linkage is bent, but also that, where steric hindrance prevents thermal agitation from disturbance of the lowest energy form, the molecule seems in this form to be dissymmetric. The result is anomalous in terms of orthodox treatment of atomic orbitals.

This behaviour is not unique; we have found apparent dissymmetries in benzoquinone, dinitrobiphenyl, dichloropiperazine and other compounds for which the chemist's intuitive common sense has dictated an orthodox symmetrical constellation of atoms. Do these heresies vitiate the conceptions of symmetry? We do not think so. Instead we consider that this evidence vitiates the use of the atomic orbital concept for any but the simplest of diatomic and perhaps some triatomic molecules. For most substances the geometry of the electron matrix cannot be defined by the atomic character of the dense matter that is embedded within this electron cloud. A new method should be devised by those who wish to amuse themselves in calculations of molecular orbitals. The field is open to the applied mathematician to define new criteria of symmetry in matter. But this time let him retain a sense of humour. The North Star is not situated so as to guide us homeward but, rather, to satisfy the electromagnetic laws.

Perhaps I go too far in criticizing intuition when it seems to go beyond the limits of good taste. I am reminded of an example where straightforward intuition might have kept us from bumbling along for twenty years, before C. K. Ingold<sup>2</sup> finally came up with the correct answer. It has long been known that certain displacement reactions occur by a basal entry of the attacking group which causes frontal ejection of the atom or group that is being displaced. The reaction:



which can be followed by radioactive tracer techniques is one of these. It is also known that the stability of the transition state through which this displacement passes is largely evaluated by the term  $B$  in the Arrhenius rate equation  $k = B e^{-E/RT}$ , which is independent of the exponential term involving temperature. Therefore dependence on  $B$  involves the hindering or helping geometry of the reacting molecule. With this knowledge it has been difficult to understand why the ease of backside approach should not vary according to the size of the atoms which might interfere with this approach. The order of reactivity of the four organic halides (Fig. 6), therefore, should be methyl > ethyl > neopentyl > *tert*.butyl, whereas the actual order considered in terms of the  $B$  factor alone is methyl > *tert*.butyl > ethyl > neopentyl. Therefore, more must be done by the attacking group than mere shoving aside of little hydrogen atoms. By calculations from experimentally determined values of bending and stretching forces

<sup>2</sup>Quarterly Reviews XI, I (1957).

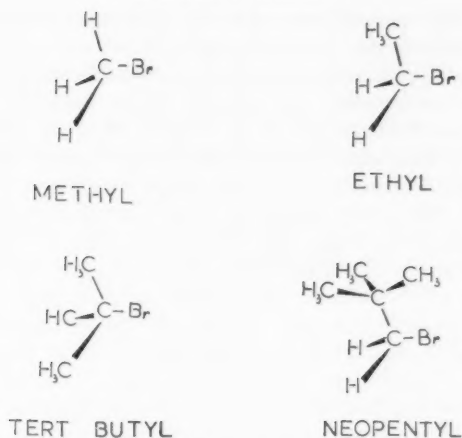


FIGURE 6

among the atoms about the central carbon atom Ingold has devised the contour maps shown in Figure 7 where the probability of the transition state (and indirectly the reaction rate) are portrayed for the temperature-dependent (enthalpy) term by the depth of the pit and the  $B$  (entropy) term by the shape of the pit in the other two directions. The vertical axis intersection with the horizontal axis "0" mark the position of the halogen

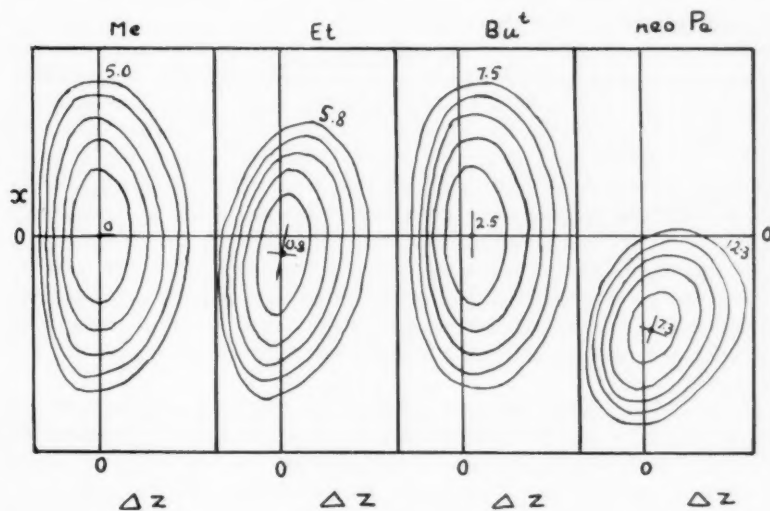


FIGURE 7

if normal rearward approach were possible and the crossed dots represent the halogen positions calculated in consideration of steric distortion. It may be seen that the symmetry of the backsides of methyl and *tert*.butyl favour substitution, but the dissymmetry of the backsides of ethyl and especially neopentyl discourage it. As Ingold correctly points out, this observation could not have been made intuitively by the organic chemists' criterion of bulk hindrance alone. However, any billiard player could assure him that the lowest energy of displacement will be achieved by central contact with the target ball. Perhaps Ingold's commentary applies to misguided intuition rather than to all intuition. Assuredly the intuitive symmetry sense of the billiard player would have helped.

During the discussions in this symposium the biologist has not been included, but now I shall try to represent him. I had not yet tried to do so because I have discussed only the chemistry of equilibrium states or of kinetic steady states. These chemistries are lifeless. But there is a chemistry of life which is very different; it is essentially dissymmetric. This statement at first seems to be anomalous when one considers the symmetrical appearance of many living things, but this apparent symmetry is incidental to the area-volume relationships required for isolated life and to the balance requirement dictated by the force of gravity; in the gross, symmetry is just as incidental to living things as it is to a volcanic cone. Actually in even the most perfect example of life some element of dissymmetry can be found, and chemistry shows this to be true in the non-intuitive sense. Indeed, life may be defined chemically as a process which defies the attainment of equilibrium or even of the steady state by application of the phenomenon of reproduction.

There are controversies over the beginning of this reproducible asymmetric process called life which it is well to avoid, largely because they are meaningless. I do not consider myself to be irreligious because I believe that one Wednesday afternoon, about three o'clock, in the age before the Pre-Cambrian, the sunlight striking the sea was reflected at the critical angle partially to become circularly polarized light. Shall we guess that by chance it was polarized dextro, and upon reflection it encountered an amino acid such as alanine which had been synthesized the night before during an electrical storm from nitrogen, oxygen and hydrogen in the atmosphere. Of course, it is doubtful that only one molecule was synthesized, and since lightning has no sense of dissymmetry, there would be equal amounts of right-handed (dextro) and left-handed (levo) varieties of alanine formed (Fig. 8).

But when circularly polarized light couples with levo alanine during a longer time interval ( $10^{-7}$  seconds) than with dextro alanine ( $10^{-8}$  seconds), because of the phase lag before re-emission of the absorbed energy, it causes the levo alanine to disintegrate. So more molecules (maybe only one!) of dextro than levo now exist in the world. Now the condition for life is started, because this one dextro molecule can react (Fig. 9) by dehydration

with an enantiomeric pair of alanine molecules to form a *dd* and a *dl* pair of diastereomers. But one of these dehydration processes is more reversible than the other; so we have a preponderance of *dd*-alanylalanine and we are on our way to a living protein. Indeed, the asymmetric growth tends to increase its specificity when the growing species becomes insoluble in the substrate and then grows in two dimensions or one dimension rather than

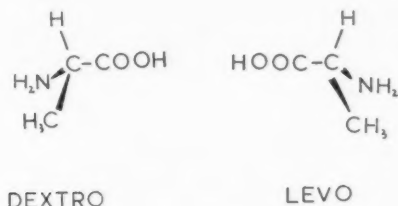


FIGURE 8

three. A description of the dissymmetric behaviour of the heterogeneous system over that of the three-dimensional homogeneous system will not be elaborated here, but it is profound.

Thus life can commence, but how can it continue? Actually an asymmetric molecule can pass its asymmetry in a chemical (that is, a biological)

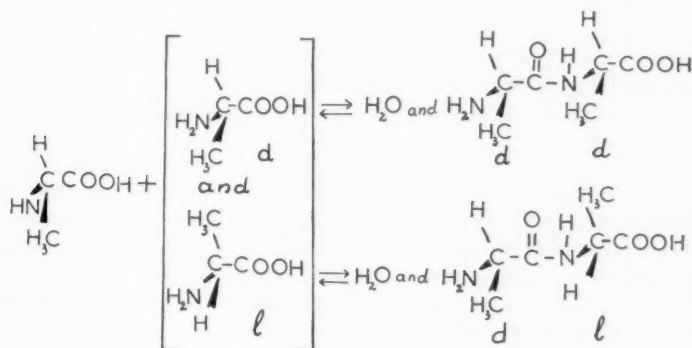


FIGURE 9

synthesis in a catalytic manner without itself being more than temporarily involved in the process. One example of this catalytic effect has been shown by Cohen and by Allentoff in the Toronto Laboratory by use of systems seemingly far removed from living processes, since water is rigidly excluded. Organomagnesium compounds ( $\text{R-Mg-X}$ ) tend, because of the tetrahedral symmetry requirement, to hold to themselves two ether ( $\text{R'OR'}$ )

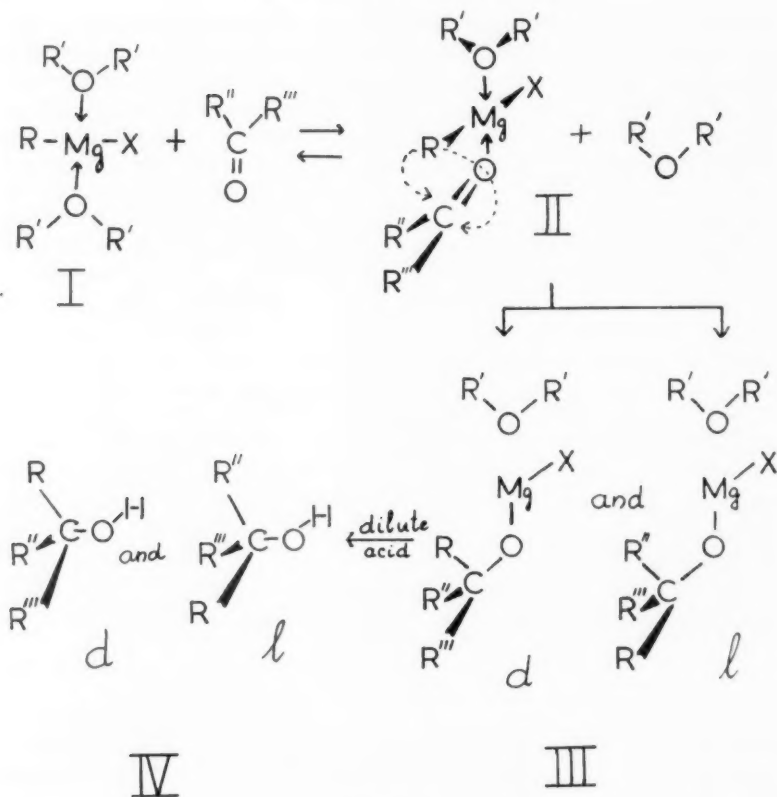


FIGURE 10

oxygen by an associative or secondary process (in Fig. 10, note structure I), but part of this ether may be displaced by a compound containing a carbonyl group in the transition state, II, from which a pair (*d* and *l*) of magnesium salts, III, are formed by rearrangement, with further loss of ether. The magnesium salts may then be converted by acidified water to give equal amounts of *d*- and *l*-alcohol, IV. But note that liberated ether may combine with another molecule of organomagnesium compound, I. Thus the ether may be said to act like a guide or a director which brings the reagents together, causes the addition to occur according to a plan, and then leaves for another similar task in the neighbouring region of the system.

Suppose now that this ether-director has an asymmetric bent. More specifically let us say that it is right-handed, as was the 2,3-dimethoxybutane which we obtained for this study from the National Research Council.

Then in the complex, II, of Figure 10 the right-handedness of the ether will favour addition on one side (→) of the carbonyl group more than on the other (←), whereas these additions would have been equally probable if the ether had been symmetrical. In consequence of its dissymmetry the N.R.C. ether will contribute its right-handed influence to the new molecule, but the ether itself will be unchanged in this process. A job of work has been done without the consumption of energy; so this job is outside the realm of thermodynamic equilibrium or steady-state chemistry. The job of work has been the transfer of asymmetry. This is the phenomenon of continuing life.

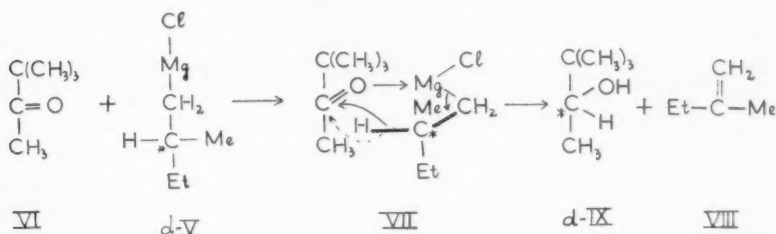


FIGURE 11

These chemical reactions point out a pattern for the genesis of the dissymmetry called life and also for its continuation. Now I want to mention the death of an asymmetric molecule as it has been described by Mosher and LaCombe (Journal of the American Chemical Society, 72 (1950), 4991) and is shown in Figure 11. A right-handed organomagnesium compound (*d-V*) is prepared from right-handed fusel oil in a symmetrical (and therefore inconsequential) ether solvent. Now a symmetrical molecule, VI, containing a carbonyl group is brought into this system. As we have seen before, the oxygen will contribute an unshared *p*-orbital electron pair to the magnesium in the tendency to accomplish the symmetry of an octet around magnesium (while the ether (not shown) may be contributing another pair). But the proximity of oxygen to magnesium will make possible the symmetry, mentioned earlier in this symposium, of the hexagonal array of atoms (VII). Within this hexagonal array electron charge may flow as I have indicated by the arrows in the picture of the transition state. During this electron flow the original dissymmetry of the organomagnesium compound disappears and this part becomes symmetrical lifeless pentene, VIII. But the dissymmetry is not lost; instead it is transferred during the shift of the hydrogen from the dying asymmetric locale to one side of the carbonyl group: (→ more than ←). Thus is born a new asymmetry (*d-IX*) from the death of the old. This is one chemist's simulation of the reproduction of life.

In conclusion it may be of interest to consider that elaborate form of life called Man. From chemical models such as I have described of things

Man sees about him he may conclude that the non-living chemistry of equilibrium or of the steady state leads inexorably to a maximum of symmetry, while by contrast the chemistry of life is essentially dissymmetric and thus transcends this law of nature. But Man would not always have drawn this conclusion. After the Renaissance it became more and more fashionable to correlate man's well-being with his adjustment to natural processes. This tendency reached its acme within the eighteenth century, but many vestiges or variations of the faith endure to the present day. Interestingly, it has been a highly religious scientist, Pierre du Nouy (*Human destiny*, New York, 1947), who has shown how the misconception arose because of a superficial knowledge of natural law. It is Du Nouy who has said that man, the highest form of life, excels himself not by being natural and conforming with the equilibria and steady states of nature, but rather by being unnatural and antagonistic toward natural law. There are many definitions of humanism but this in large part is my definition.



## Differentiability Properties of Arcs of Order $n + 1$ in Conformal $n$ -space

N. D. LANE

Presented by R. L. JEFFERY, F.R.S.C.

**Introduction.** In (6, §3.2 and §3.5) it was shown that the end-points of arcs of order three in the conformal plane not only satisfy automatically the conditions for conform-differentiability given in (5, §5 and §7), but that they also satisfy a stronger set of conditions (6, §3.1). In the present paper, the author shows that an end-point of an arc of order  $n + 1$  in conformal  $n$ -space is automatically strongly differentiable. This can be done directly, as in §4.6, by embedding conformal  $n$ -space in projective  $(n + 1)$ -space. It is, however, of some interest to discuss this problem using only conformal methods. In §§4.2–4.5 most of the differentiability properties are proved without resort to central projection. To this end, the multiplicities with which the tangent  $m$ -spheres meet this arc at its end-point are also discussed in §3.

### 1. Notation

**1.1** The letters  $p, t, P, \dots$  denote points in real conformal  $n$ -space;  $S^{(m)}$  will denote an  $m$ -sphere. When there is no ambiguity, the superscript  $(n - 1)$  will be omitted in the case of  $S^{(n-1)}$ ; thus an  $(n - 1)$ -sphere will usually be denoted by  $S$  alone. Such an  $(n - 1)$ -sphere  $S$  decomposes the  $n$ -space into two open regions, its *interior*  $\tilde{S}$  and its *exterior*  $\bar{S}$  (cf. 3, §1). If  $S$  is a point  $(n - 1)$ -sphere, we may assume that  $\tilde{S}$  is void.

**1.2** The definitions of convergence, arc, end-point, interior point, neighbourhood, support and intersection, are identical with those given in (4, §1).

**1.3** Let  $p$  be a fixed point of an arc  $A$  and let  $t$  be a variable point of  $A$ . Let  $1 \leq m < n$ . If  $p, P_1, \dots, P_{m+1}$  do not lie on the same  $(m - 1)$ -sphere, then there exists a unique  $m$ -sphere  $S^{(m)}(P_1, \dots, P_{m+1}, p)$  through these points. It is convenient to denote this  $m$ -sphere by

$$S^{(m)} = S^{(m)}(p, P_1, \dots, P_{m+1}, \tau_0);$$

here  $\tau_0$  indicates that this  $m$ -sphere passes through  $p$ . In the following, the  $m$ -sphere  $S^{(m)}(P_1, \dots, P_{m+1}, \tau_r)$  is defined inductively by means of the

conditions  $\Gamma_r^{(m)}$  given below; (the  $\tau_r$  in the symbol  $S^{(m)}(P_1, \dots, P_{m+1-r}; \tau_r)$  indicates that this sphere is a *tangent  $m$ -sphere* of the arc  $A$  at the point  $p$  meeting  $A$   $r+1$  times at  $p$ ). We call  $A$   $(m+1)$  times differentiable at  $p$  if the following sequence of conditions is satisfied (3, §4).

$\Gamma_r^{(m)}$ , ( $r = 1, 2, \dots, m+1$ ): if the parameter  $t$  is sufficiently close to, but different from, the parameter  $p$ , then the  $m$ -sphere  $S_r^{(m)}(P_1, \dots, P_{m+1-r}, t; \tau_{r-1})$  is uniquely defined. It converges if  $t$  tends to  $p$ . Thus its limit sphere, which will be denoted by  $S^{(m)}(P_1, \dots, P_{m+1-r}; \tau_r)$ , will be independent of the way  $t$  converges to  $p$  [condition  $\Gamma_{m+1}^{(m)}$  reads:  $S^{(m)}(t; \tau_m)$  exists and converges to  $S_{m+1}^{(m)} = S^{(m)}(\tau_{m+1})$ ].

$A$  is called *once differentiable* if  $\Gamma_1^{(1)}$  is satisfied. The point  $p$  is called a *differentiable point* of  $A$  if  $A$  is  $n$  times differentiable at  $p$ .

$\tau_r^{(m)}$  will denote the family of all the  $S_r^{(m)}$ 's ( $\tau_r$  will mean  $\tau_r^{(n-1)}$ ). In particular,  $\tau_{m+1}^{(m)}$  consists only of  $S_{m+1}^{(m)}$ , the *osculating  $m$ -sphere* of  $A$  at  $p$ .  $S_0^{(0)}$  denotes a pair of points  $P, p$  and  $S_1^{(0)}$  denotes  $p$ .

Differentiability implies, in particular, that:

- (i)  $\tau_0^{(m)} \supset \tau_1^{(m)} \supset \dots \supset \tau_{m+1}^{(m)}$  (cf. 3, Theorem 3, Corollary 4).
- (ii)  $\tau_r^{(m)}$  is satisfied for all  $m$  and  $r$  such that  $1 \leq m < n$  and  $1 \leq r \leq m+1$  (cf. 3, Theorem 2).
- (iii)  $S^{(m-1)}(P_1, \dots, P_{m-r}; \tau_r) = \bigcap_P S^{(m)}(P_1, \dots, P_{m-r}, P; \tau_r),$   
( $1 < m < n; 1 \leq r \leq m$ ).

Also if

$$P \not\subset S_r^{(m-1)} = S^{(m-1)}(P_1, \dots, P_{m-r}; \tau_r),$$

then

$$S^{(m)}[P, S_r^{(m-1)}] = S^{(m)}(P_1, \dots, P_{m-r}, P; \tau_r),$$

(cf. 3, Theorem 2).

(iv) If  $S_r^{(r-1)} \neq p$ , then  $\tau_r^{(m)}$  consists of all the  $m$ -spheres through  $S_r^{(r-1)}$ . If  $S_r^{(r-1)} = p$ , then  $\tau_r^{(m)}$  is the set of all the  $m$ -spheres which touch any  $S_r^{(r)} \subset \tau_r^{(r)}$  at  $p$  (3, Theorem 3). There is only one  $m$ -sphere of  $\tau_r^{(m)}$  through  $m+1-r$  points which do not lie on the same  $S_r^{(m-1)}$  (cf. 3, Theorem 3, Corollary 2).

**1.4** An arc  $A$  is said to be of *finite order* if  $A$  has only a finite number of points in common with any  $(n-1)$ -sphere. If the least upper bound of these numbers is finite, then it is called the *conformal order* of  $A$  and  $A$  is said to be of *bounded order*.

## 2. Differentiability at an end-point of an arc of finite order

The proof given in (6, §3.2), that an end-point of an arc of finite order in the conformal plane is automatically differentiable can be readily extended to  $n$ -dimensions (2).

**THEOREM 1.** *Let  $p$  be an end-point of an arc  $A$  of finite order. Then  $A$  is automatically differentiable at  $p$ .*

*Proof.* Suppose  $r$  is the smallest integer for which  $\Gamma_r^{(n-1)}$  is not satisfied. Then for some set of points  $P_1, \dots, P_{n-r}$  such that  $P_1, \dots, P_{n-r}, p$  do not lie on the same  $(n-r-2)$ -sphere, there are two sequences of points  $t_{2k}$  and  $t_{2k+1}$ , different from  $p$  and convergent on  $A$  to  $p$ , such that the  $(n-1)$ -spheres

$$S_{2k} = S(P_1, \dots, P_{n-r}, t_{2k}; \tau_{r-1}), S_{2k+1} = S(P_1, \dots, P_{n-r}, t_{2k+1}; \tau_{r-1})$$

converge to different limit spheres  $S_0$  and  $S_1$  respectively. We may assume that  $t_{r+1}$  lies between  $p$  and  $t_r$ . If  $k$  is large,  $S_{2k}$  ( $S_{2k+1}$ ) will lie close to  $S_0$  ( $S_1$ ). Let  $S$  and  $S'$  be two  $(n-1)$ -spheres through  $S^{(n-2)}(P_1, \dots, P_{n-r}; \tau_{r-1})$  which separate  $S_r$  and  $S_{r+1}$ . (In the case  $r = n$  and  $S_{n-1}^{(n-2)} = p$ ,  $S_r = S(t_r; \tau_{n-1})$  and  $S_{r+1} = S(t_{r+1}; \tau_{n-1})$  touch at  $p$  and we may take  $S' = p$ .) Then  $S \cup S'$  will separate  $S_r$  and  $S_{r+1}$  and therefore also  $t_r$  and  $t_{r+1}$  for every large  $r$ . Hence the subarc of  $A$  bounded by  $t_r$  and  $t_{r+1}$  will meet  $S \cup S'$  in at least one point. Thus  $A$  will meet  $S \cup S'$  infinitely often. Since, however,  $A$  has finite order,  $\Gamma_r^{(n-1)}$  must be satisfied.

It may be remarked that actually only the cases  $r = 1$  and  $r = n$  need to be proved above since it is shown in (3, Theorem 4), that  $\Gamma_1^{(n-1)}$  implies  $\Gamma_r^{(n-1)}$  for  $r = 2, 3, \dots, n-1$ , and even for  $r = n$  if  $S_{n-1}^{(n-2)} \neq p$ .

### 3. Multiplicities at an end-point of $A_{n+1}$

**3.1** Let  $A_{n+1}$  denote an open arc of order  $n+1$ . It is clear that a point of  $A_{n+1}$  converges if its parameter tends to one of the end-points of the parameter interval. Thus  $A_{n+1}$  has two well-defined end-points. Let  $p$  be one of them. We introduce multiplicities at  $p$  such that  $p$  is counted  $r+1$  times on any  $m$ -sphere of  $\tau_r^{(m)} - \tau_{r+1}^{(m)}$  and a point of support is counted at least twice on  $A_{n+1}$ . We wish to prove

**THEOREM 2.** *No  $m$ -sphere meets  $A_{n+1} \cup p$  more than  $m+2$  times, that is, the inclusion of  $p$  and the introduction of multiplicities does not alter the order of  $A_{n+1}$ .*

**3.2** Suppose an  $(n-1)$ -sphere  $S$  meets  $A_{n+1}$  at  $t_1, t_2, \dots, t_n$  and  $t$ . Then  $t \notin S^{(n-2)}(t_1, \dots, t_n)$ . If  $t$  is a point of support, there is an  $(n-1)$ -sphere close to  $S$ , through  $S^{(n-2)}(t_1, \dots, t_n)$ , which intersects  $A_{n+1}$  at least twice near  $t$ . Thus this sphere meets  $A_{n+1}$  at least  $n+2$  times. Hence:

*An  $(n-1)$ -sphere through  $(n+1)$  points of  $A_{n+1}$  intersects  $A_{n+1}$  at each of them.*

Suppose now that  $p \subset S$ . Now  $p$  does not lie on both  $S^{(n-2)}(t_1, \dots, t_n)$  and  $S^{(n-2)}(t_2, \dots, t_n, t)$  otherwise  $t_1, \dots, t_n, t$  would lie on the same  $S^{(n-2)}$  through  $p$ . Suppose, for example, that  $p \notin S^{(n-2)}(t_1, \dots, t_n)$ . Then  $S = S[p, S^{(n-2)}(t_1, \dots, t_n)]$ . Choose disjoint neighbourhoods  $B$  of  $t$  and  $M$  of  $p$  on  $A$  which do not include  $t_1, \dots, t_n$ . If  $u$  converges in  $M$  to  $p$  then

$S' = S[u, S^{(n-2)}(t_1, \dots, t_n)]$  converges to  $S$  and hence  $S'$  will intersect  $B$  if  $u$  is sufficiently close to  $p$ . Thus  $S'$  will meet  $A_{n+1}$  in not less than  $n+2$  points. Hence:

*An  $(n-1)$ -sphere through  $p$  and  $n$  points of  $A_{n+1}$  does not meet  $A_{n+1}$  elsewhere.*

Suppose an  $(n-1)$ -sphere  $S$  through  $p, t_1, \dots, t_n$  supports  $A_{n+1}$  at  $t_n$ . From the above,  $t_n \notin S^{(n-2)} = S^{(n-2)}(p, t_1, \dots, t_{n-1})$  and a suitable  $(n-1)$ -sphere  $S'$  through  $S^{(n-2)}$  close to  $S$  will intersect  $A_{n+1}$  twice near  $t_n$ . Thus:

*An  $(n-1)$ -sphere through  $p$  and  $n$  points of  $A_{n+1}$  intersects  $A_{n+1}$  at each of these points.*

The above remarks can readily be extended as follows:

*If an  $(n-1)$ -sphere  $S^{(n-1)}(P_1, \dots, P_k, t_1, \dots, t_{n-2k+1}, u_1, \dots, u_k)$  supports  $A_{n+1}$  at  $u_1, \dots, u_k$  and intersects  $A_{n+1}$  at  $t_1, \dots, t_{n-2k+1}$  then it does not meet  $A_{n+1}$  again.*

*Proof.* From the above, the statement is true if  $k=0$ . Assume that it holds for all values  $\leq k$ , where  $0 \leq 2k < n$ . Suppose an  $(n-1)$ -sphere

$$S = S^{(n-1)}(P_1, \dots, P_k, t_1, \dots, t_{n-2k}, u_1, \dots, u_{k+1})$$

intersects  $A_{n+1}$  at the points  $t_i$  and supports at points  $u_j$ . Choose disjoint neighbourhoods  $M_i$  of the  $t_i$  and  $N_j$  of the  $u_j$ . Now,

$$S \supset S^{(n-2)} = S^{(n-2)}(P_1, \dots, P_k, t_1, \dots, t_{n-2k}, u_1, \dots, u_k)$$

and  $S = S^{(n-1)}[u_{k+1}; S^{(n-2)}]$  if  $u_{k+1} \notin S^{(n-2)}$ . In this case, a suitable  $(n-1)$ -sphere through  $S^{(n-2)}$  and close to  $S$  will meet each  $M_i$  and also will meet each  $N_j$  at least twice (counting any points of support twice). Thus this  $(n-1)$ -sphere will meet  $A_{n+1}$  more than  $n+1$  times. Essentially, there is no loss in generality in assuming that  $u_{k+1} \notin S^{(n-2)}$ . For, suppose  $u_{k+1} \in S^{(n-2)}$ . Now

$$u_{k+1} \notin S^{(n-k-2)}(t_1, \dots, t_{n-2k}, u_1, \dots, u_k).$$

Choose  $Q_\lambda \subset S$ ,  $\lambda = 1, 2, \dots, k$ , in turn, such that

$$Q_\lambda \not\subset S^{(n-k-3+\lambda)}(Q_1, \dots, Q_{\lambda-1}, t_1, \dots, t_{n-2k}, u_1, \dots, u_k)$$

and

$$u_{k+1} \not\subset S^{(n-k-2+\lambda)}(Q_1, \dots, Q_\lambda, t_1, \dots, t_{n-2k}, u_1, \dots, u_k).$$

Thus, for  $\lambda = k$ ,

$$S = S^{(n-1)}[S^{(n-2)}(Q_1, \dots, Q_k, t_1, \dots, t_{n-2k}, u_1, \dots, u_k); u_{k+1}].$$

This is a special case of Lemma 5, below, if in this Lemma we set  $r = -1$  and omit the symbol  $\tau_r$  throughout.

Let  $0 \leq r < n$ . Suppose that for every choice of  $t_1, \dots, t_{n-r}$  on  $A_{n+1}$ , the  $(n-1)$ -sphere  $S^{(n-1)}(t_1, \dots, t_{n-r}; \tau_r)$  does not meet  $A_{n+1}$  again. If  $S_{r+1}^{(n-1)} = S^{(n-1)}(t_1, \dots, t_{n-r-1}; \tau_{r+1})$  intersects  $A_{n+1}$  at a point  $u \neq t_i$ , then for every  $t$  sufficiently close to  $p$  on  $A_{n+1}$ ,  $S^{(n-1)}(t_1, \dots, t_{n-r-1}, t; \tau_r)$  will intersect  $A_{n+1}$  again near  $u$ . This yields:

LEMMA 1.  $S^{(n-1)}(t_1, \dots, t_{n-r}; \tau_r)$  does not intersect  $A_{n+1}$  at another point ( $r = 0, 1, \dots, n$ ).

If  $S_{r+1}^{(n-1)} = S^{(n-1)}(t_1, \dots, t_{n-r-1}; \tau_{r+1})$  supports  $A_{n+1}$  at another point  $u$ , then

$$u \not\subset S_r^{(n-2)} = S^{(n-2)}(t_1, \dots, t_{n-r-1}; \tau_r),$$

otherwise if  $v \in A_{n+1}$ ,  $v \not\subset S_r^{(n-2)}$ , then

$$S^{(n-1)}[v, S_r^{(n-2)}] = S^{(n-1)}(t_1, \dots, t_{n-r-1}, v; \tau_r)$$

will meet  $A_{n+1}$  in more than  $n - r$  points. Thus a suitable  $(n - 1)$ -sphere through  $S_r^{(n-2)}$ , close to  $S_{r+1}^{(n-1)}$ , will be an  $S_r^{(n-1)}$  which meets  $A_{n+1}$  in more than  $n - r$  points.

Hence:

LEMMA 2.  $S^{(n-1)}(t_1, \dots, t_{n-r}; \tau_r)$  does not support  $A_{n+1}$  at another point ( $r = 0, 1, \dots, n$ ).

The following results can now be readily proved:

LEMMA 3.  $S^{(n-1)}(t_1, \dots, t_{n-r}; \tau_r)$  intersects  $A_{n+1}$  at  $t_1, \dots, t_{n-r}$ .

*Proof.* Lemmas 1 and 2 imply that  $t_1 \not\subset S^{(n-2)}(t_2, \dots, t_{n-r}; \tau_r)$ . Thus

$$S_r^{(n-1)} = S^{(n-1)}(t_1, \dots, t_{n-r}; \tau_r) = S^{(n-1)}[t_1, S^{(n-2)}(t_2, \dots, t_{n-r}; \tau_r)].$$

If  $S_r^{(n-1)}$  supports  $A_{n+1}$  at  $t_1$ , then a suitable  $(n - 1)$ -sphere through  $S^{(n-2)}(t_2, \dots, t_{n-r}; \tau_r)$  and close to  $S_r^{(n-1)}$  will intersect  $A_{n+1}$  at least twice near  $t_1$ . Altogether this new  $(n - 1)$ -sphere of  $\tau_r$  will meet  $A_{n+1}$  in at least  $n - r + 1$  points.

LEMMA 4.  $S^{(m)}(t_1, \dots, t_{m+1-r}; \tau_r)$  does not meet  $A_{n+1}$  again.

*Proof.* Suppose  $S^{(m)}(t_1, \dots, t_{m+1-r}; \tau_r)$  meets  $A_{n+1}$  again at  $t$ . Choose  $u_1, \dots, u_{n-m-1}$  in turn,  $\neq t$ , on  $A_{n+1}$  so that

$$u_{k+1} \not\subset S^{(m+k)}(t_1, \dots, t_{m+1-r}, u_1, \dots, u_k; \tau_r), \quad 0 \leq k < n - m - 1.$$

Then  $S^{(n-1)}(t_1, \dots, t_{m+1-r}, u_1, \dots, u_{n-m-1}; \tau_r)$  will meet  $A_{n+1}$  in more than  $n - r$  points.

LEMMA 5. If  $S^{(n-1)}(P_1, \dots, P_k, t_1, \dots, t_{n-r-2k}, u_1, \dots, u_k; \tau_r)$  supports  $A_{n+1}$  at  $u_1, \dots, u_k$  and intersects  $A_{n+1}$  at  $t_1, \dots, t_{n-r-2k}$  then it does not meet  $A_{n+1}$  again.

*Proof.* By Lemmas 1 and 2, the statement is true if  $k = 0$ . Assume that it holds for all values  $\leq k$ , where  $0 \leq 2k < n - r$ . Suppose an  $(n - 1)$ -sphere

$$S_r = S^{(n-1)}(P_1, \dots, P_k, t_1, \dots, t_{n-r-2k-1}, u_1, \dots, u_{k+1}; \tau_r)$$

intersects  $A_{n+1}$  at the points  $t_i$  and supports at points  $u_j$ . Choose disjoint neighbourhoods  $M_i$  of the  $t_i$  and  $N_j$  of the  $u_j$ . Now,

$$S_r \supset S_r^{(n-2)} = S_r^{(n-2)}(P_1, \dots, P_k, t_1, \dots, t_{n-r-2k-1}, u_1, \dots, u_k; \tau_r)$$

and  $S_r = S^{(n-1)}[u_{k+1}; S_r^{(n-2)}]$  if  $u_{k+1} \not\subset S_r^{(n-2)}$  and  $S_r^{(n-2)} \neq p$ . In this case, a suitable  $(n-1)$ -sphere through  $S_r^{(n-2)}$  and close to  $S_r$  will be an  $(n-1)$ -sphere of  $\tau_r$  which meets each  $M_i$  and also meets each  $N_j$  at least twice (counting any points of support twice). If  $S_r^{(n-2)} = p$ , a similar argument based on §1.3, (iv), can be used. Thus this  $(n-1)$ -sphere of  $\tau_r$  will meet  $A_{n+1}$  more than  $n-r$  times. Essentially, there is no loss in generality in assuming that  $u_{k+1} \not\subset S^{(n-2)}$ . For, suppose  $u_{k+1} \subset S_r^{(n-2)}$ . Now

$$u_{k+1} \not\subset S^{(n-k-2)}(t_1, \dots, t_{n-r-2k-1}, u_1, \dots, u_k; \tau_r).$$

Choose  $Q_\lambda \subset S_r$ ,  $\lambda = 1, 2, \dots, k$ , in turn, such that

$$Q_\lambda \not\subset S^{(n-k-3+\lambda)}(Q_1, \dots, Q_{\lambda-1}, t_1, \dots, t_{n-r-2k-1}, u_1, \dots, u_k; \tau_r)$$

and

$$u_{k+1} \not\subset S^{(n-k-2+\lambda)}(Q_1, \dots, Q_\lambda, t_1, \dots, t_{n-r-2k-1}, u_1, \dots, u_k; \tau_r).$$

Thus

$$S_r = S^{(n-1)}[S^{(n-2)}(Q_1, \dots, Q_k, t_1, \dots, t_{n-r-2k-1}, u_1, \dots, u_k; \tau_r); u_{k+1}].$$

#### 4. Strong differentiability

**4.1** An arc  $A$  will be called *strongly differentiable* at  $p$  if the  $m$ -sphere  $S^{(m)}(P_1, \dots, P_{m+1-r}, t_1, \dots, t_j; \tau_{r-j})$  converges to  $S^{(m)}(P_1, \dots, P_{m+1-r}; \tau_r)$  whenever  $t_1, \dots, t_j$  converge on  $A$  to  $p$  ( $j = 1, 2, \dots, r+1$ ).

**THEOREM 3.** *Let  $p$  be an end-point of an arc  $A_{n+1}$  of order  $n+1$ . Then  $A_{n+1}$  is automatically strongly differentiable at  $p$ .*

While the required conditions, with one exceptional case, for strong differentiability are established conformally in §§4.2-4.5, a short and independent proof of Theorem 3 is given in §4.6.

**4.2** Let  $B$  be an open subarc of  $A_{n+1}$  bounded by  $p$  and any point  $e$  of  $A_{n+1}$ . Let  $d$  be any point of  $A_{n+1}$  outside  $B \cup e$ . The  $(n-1)$ -sphere  $S$  with  $d \not\subset S$  will be oriented such that  $d \subset \tilde{S}$ . The set of all these  $(n-1)$ -spheres contains all the  $(n-1)$ -spheres which meet  $A_{n+1} \cup p$   $n+1$  times in  $p \cup B \cup e$ . Their orientation is continuous. In particular, the region  $\tilde{S}^{(n-1)}(u_1, \dots, u_{n-r}; \tau_r)$  depends continuously on  $u_1, \dots, u_{n-r}$  when these mutually distinct points range through  $p \cup B \cup e$ .

In the following,  $u_1, \dots, u_{n-r}, t_1, \dots, t_j, p$  are assumed to be mutually distinct and to lie on  $B \cup e$  in the indicated order. Thus  $d \subset \tilde{S}_r^{(n-1)}(u_1, \dots, u_{n-r}; \tau_r)$  and, on account of Theorem 2, a point  $v$  between  $u_i$  and  $u_{i+1}$  on  $B$  will lie in  $\tilde{S}_r^{(n-1)}$  or in  $\tilde{S}_r^{(n-1)}$  according as  $i$  is even or odd.

**4.3** Let  $u_1, \dots, u_{n-r}$  be fixed points on  $B$ . Let  $1 \leq r \leq n$ . It will first be proved that as  $t_1, \dots, t_j$  tend to  $p$  on  $B$  the condition  $P_{rj}: S^{(n-1)}(u_1, \dots, u_{n-r}, t_1, \dots, t_j; \tau_{r-j}) \rightarrow S^{(n-1)}(u_1, \dots, u_{n-r}; \tau_r)$  holds ( $j = 1, 2, \dots, r+1$ ).

*Proof* (by induction with respect to  $r$  and  $j$ ). Assume that for each  $k$  such that  $1 \leq k \leq r$ , and for each  $j$  such that  $1 \leq j \leq k+1$ , the proposition  $P_{kj}$  holds. Theorem 1 implies that  $P_{r1}$  holds. On the assumption that  $P_{rj}$  holds, it will be proved that  $P_{r,j+1}$  also holds ( $1 \leq j \leq r$ ).

Since  $t_{j+1}$  lies either in the region

$$\tilde{S}_{r-j} \cap \tilde{S}_{r-j-1} = \tilde{S}(u_1, \dots, u_{n-r}, t_1, \dots, t_j; \tau_{r-j}) \cap \tilde{S}(u_1, \dots, u_{n-r}, u_{n-r+1}, t_1, \dots, t_j; \tau_{r-j-1}),$$

or else in the region  $\tilde{S}_{r-j} \cap \tilde{S}_{r-j-1}$ , it follows that  $S(u_1, \dots, u_{n-r}, t_1, \dots, t_{j+1}; \tau_{r-j-1}) \subset (\tilde{S}_{r-j} \cap \tilde{S}_{r-j-1}) \cup (\tilde{S}_{r-j} \cap \tilde{S}_{r-j-1}) \cup (S_{r-j} \cap S_{r-j-1})$ .

Thus any limit sphere  $S$  of  $S(u_1, \dots, u_{n-r}, t_1, \dots, t_{j+1}; \tau_{r-j-1})$  lies in  $[\tilde{S}(u_1, \dots, u_{n-r}; \tau_r) \cap \tilde{S}(u_1, \dots, u_{n-r+1}; \tau_{r-1})] \cup [\tilde{S}(u_1, \dots, u_{n-r}; \tau_r) \cap \tilde{S}(u_1, \dots, u_{n-r+1}; \tau_{r-1})] \cup$

$$S(u_1, \dots, u_{n-r}; \tau_r) \cup S(u_1, \dots, u_{n-r+1}; \tau_{r-1}).$$

This holds for every choice of  $u_{n-r+1}$  on  $B$  while  $S$  is independent of  $u_{n-r+1}$ . Letting  $u_{n-r+1}$  tend to  $p$ , we obtain  $S \subset S(u_1, \dots, u_{n-r}; \tau_r)$ . If  $r < n$ , this implies that  $S = S(u_1, \dots, u_{n-r}; \tau_r)$ . Let  $r = n$ .

If  $S_{n-1}^{(n-2)} \neq p$ , then  $S = S_n^{(n-1)}$ . If  $S_n^{(n-1)} = p$ , then  $S = p$  also. Suppose  $S_{n-1}^{(n-2)} = p$ , while  $S_n^{(n-1)} \neq p$ . Then the  $(n-1)$ -spheres of  $\tau_{n-1}$  all touch at  $p$  and  $S(t_1, \dots, t_{j+1}; \tau_{n-j-1})$  separates the regions

$$\tilde{S}_{n-j} \cap \tilde{S}_{n-j-1} = \tilde{S}(t_1, \dots, t_j; \tau_{n-j}) \cap \tilde{S}(u_1, t_1, \dots, t_j; \tau_{n-j-1}), \tilde{S}_{n-j} \cap \tilde{S}_{n-j-1}.$$

Therefore  $S$  separates the regions

$$\tilde{S}(\tau_n) \cap \tilde{S}(u_1; \tau_{n-1}) = \tilde{S}(\tau_n), \tilde{S}(\tau_n) \cap \tilde{S}(u_1; \tau_{n-1}) = \tilde{S}(u_1; \tau_{n-1}).$$

Thus  $S \neq p$  and hence  $S = S_n^{(n-1)}$ .

**4.4** Let  $0 \leq r \leq m+1 \leq n$ ,  $1 \leq j \leq r+1$ . Then as  $t_1, \dots, t_j$  tend to  $p$  the proposition

$$P_{rj}^{(m)}: S^{(m)}(u_1, \dots, u_{m+1-r}, t_1, \dots, t_j; \tau_{r-j}) \rightarrow S^{(m)}(u_1, \dots, u_{m+1-r}; \tau_r)$$

holds.

*Proof* (by induction with respect to  $m$ ). The statement is valid if  $m = n-1$ , by §4.3. It will be shown that  $P_{rj}^{(m)}$  implies  $P_{rj}^{(m-1)}$  for  $r \leq m \leq n$ . The assumption

$$S^{(m)}(u_1, \dots, u_{m+1-r}, t_1, \dots, t_j; \tau_{r-j}) \rightarrow S^{(m)}(u_1, \dots, u_{m+1-r}; \tau_r)$$

implies

$$\bigcap S^{(m)}(u_1, \dots, u_{m+1-r}, t_1, \dots, t_j; \tau_{r-j}) \rightarrow \bigcap_{u_{m+1-r}} S^{(m)}(u_1, \dots, u_{m+1-r}; \tau_r).$$

Cf.  $u_{m+1-r}$ .

Thus by §1.3, (iii),

$$S^{(m-1)}(u_1, \dots, u_{m-r}, t_1, \dots, t_j; \tau_{r-j}) \rightarrow S^{(m-1)}(u_1, \dots, u_{m-r}; \tau_r).$$



In particular, this yields

$$S^{(r-1)}(t_1, \dots, t_j; \tau_{r-j}) \rightarrow S^{(r-1)}(\tau_r).$$

**4.5** Let  $1 \leq m < n$ ,  $1 \leq r \leq m$ ,  $1 \leq j \leq r+1$ . It will next be proved that as  $t_1, \dots, t_j$  tend to  $p$  on  $A_{n+1}$

$$S^{(m)}(P_1, \dots, P_{m+1-r}, t_1, \dots, t_j; \tau_{r-j}) \rightarrow S^{(m)}(P_1, \dots, P_{m+1-r}; \tau_r),$$

whenever  $S^{(m-1)}(P_1, \dots, P_{m+1-r}; \tau_{r-1}) \not\subset \tau_r^{(m-1)}$ .

*Proof.* Let  $S^{(m)}$  be any limit sphere of

$$S_{r-j}^{(m)} = S^{(m)}(P_1, \dots, P_{m+1-r}, t_1, \dots, t_j; \tau_{r-j}).$$

Since  $S_{r-j}^{(m)} \supset S^{(r-1)}(t_1, \dots, t_j; \tau_{r-j})$ , it follows that  $S^{(m)} \supset S_r^{(r-1)}$ . If  $S_r^{(r-1)} \neq p$ , then  $S^{(m)}$  is the unique  $m$ -sphere of  $\tau_r^{(m)}$  through  $P_1, \dots, P_{m+1-r}$  (cf. §1.3 (iv)). Let  $S_r^{(r-1)} = p$ . Suppose  $u$  is a fixed point on  $B$ . As  $t_1, \dots, t_j$  tend to  $p$ ,

$$\lim \angle \{S^{(r)}[P_1, S^{(r-1)}(t_1, \dots, t_j; \tau_{r-1})], S^{(r)}[u, S^{(r-1)}(t_1, \dots, t_j; \tau_{r-j})]\} \rightarrow 0,$$

(3, Theorem 3). Thus  $S^{(r)}[P_1, S^{(r-1)}(t_1, \dots, t_j; \tau_{r-j})]$  converges to the unique  $r$ -sphere through  $P_1$  which touches  $S^{(r)}(u; \tau_r)$  at  $p$ , that is,

$$\lim S^{(r)}[P_1, S^{(r-1)}(t_1, \dots, t_j; \tau_{r-j})] = S^{(r)}(P_1; \tau_r).$$

By §1.3, (iv),  $S^{(m)}$  is the unique  $m$ -sphere of  $\tau_r^{(m)}$  through  $P_1, \dots, P_{m+1-r}$ , that is,  $S^{(m)} = S^{(m)}(P_1, \dots, P_{m+1-r}; \tau_r)$ .

**4.6** If  $S^{(m-1)}(P_1, \dots, P_{m+1-r}; \tau_{r-1}) \subset \tau_r^{(m-1)}$ , the method of §4.5 fails to show that  $S^{(m)}(P_1, \dots, P_{m+1-r}, t_1, \dots, t_j; \tau_{r-j}) \rightarrow S^{(m)}(P_1, \dots, P_{m+1-r}; \tau_r)$ . The whole of Theorem 3, however, can be proved, without assuming the results of §§4.2-4.5, by representing conformal  $n$ -space on an  $n$ -sphere in projective  $(n+1)$ -space and making use of central projection.

It is proved in (7) that an end-point of an arc  $A_n$  of linear order  $n$  in projective  $n$ -space is strongly differentiable. In particular, the linear  $r$ -space ( $r = 1, 2, \dots, n-1$ ) through  $r+1$  mutually distinct points  $t_1, \dots, t_{r+1}$  of  $A_n$  converges, as these points converge to an end-point  $p$  of  $A_n$ , to a unique limiting linear osculating  $r$ -space, which will be denoted by  $L_r^{(n)}(p)$ . Let  $r \leq m < n$ . If  $P_1, \dots, P_{m-r+1}, p$  are fixed independent points it is easy to prove by induction with respect to  $m$  that the linear  $m$ -space through  $P_1, \dots, P_{m-r+1}, t_1, \dots, t_r$  converges to a unique  $m$ -space through  $L_r^{(n)}(p)$  and  $P_1, \dots, P_{m-r+1}$ .

*Proof.* A central projection, with  $P_{m-r+1}$  as centre, on an  $(n-1)$ -space which does not contain  $P_{m-r+1}$  will map the arc  $A_n$  of linear order  $n$  on an arc  $A'_n$  of order  $n$ . Let  $P'_1, \dots, P'_{m-r}, t'_1, \dots, t'_r$  be the images of  $P_1, \dots, P_{m-r}, t_1, \dots, t_r$ . It is proved in (1) that an arc of linear order  $n$  in projective  $(n-1)$ -space is the union of a bounded number of arcs of order

$n - 1$ . Hence if the  $t_i$  are sufficiently close to  $p$ , the  $t'_i$  will belong to a sub-arc  $A'_{n-1}$  of order  $n - 1$  of  $A'_n$ . Suppose that  $P'_1, \dots, P'_{m-r}, t'_1, \dots, t'_r$  define a unique linear  $(m - 1)$ -space which converges uniquely as the  $t'_i$  converge on  $A'_{n-1}$  to  $p$ . In the case  $m = r$ , this is automatically true by the result (cf. 7) mentioned above. It then follows that  $P_1, \dots, P_{m-r+1}, t_1, \dots, t_r$  define a unique linear  $m$ -space which converges uniquely as  $t_1, \dots, t_r$  converge to  $p$ .

Now, conformal  $n$ -space can be represented on an  $n$ -sphere in projective  $(n + 1)$ -space. Each  $(n - 1)$ -sphere  $S^{(n-1)}$  on this  $n$ -sphere will define a linear  $n$ -space (namely, that linear  $n$ -space which contains  $S^{(n-1)}$ ). In particular, an arc of conformal order  $n + 1$  on the  $n$ -sphere will therefore become a (spherical) arc of linear order  $n + 1$  in projective  $(n + 1)$ -space. By the above,  $S^{(m)}(P_1, \dots, P_{m+1-r}, t_1, \dots, t_j; \tau_{r-j})$  may be associated with an  $(r - j)$ -fold tangent linear  $m$ -space through  $P_1, \dots, P_{m+1-r}$  which converges uniquely as the  $t_i \rightarrow p$  on  $A_{n+1}$ . The intersection of this linear  $m$ -space with the  $n$ -sphere determines a unique limiting  $m$ -sphere which, by §1.3, (iv), coincides with  $S^{(m)}(P_1, \dots, P_{m+1-r}; \tau_r)$ .

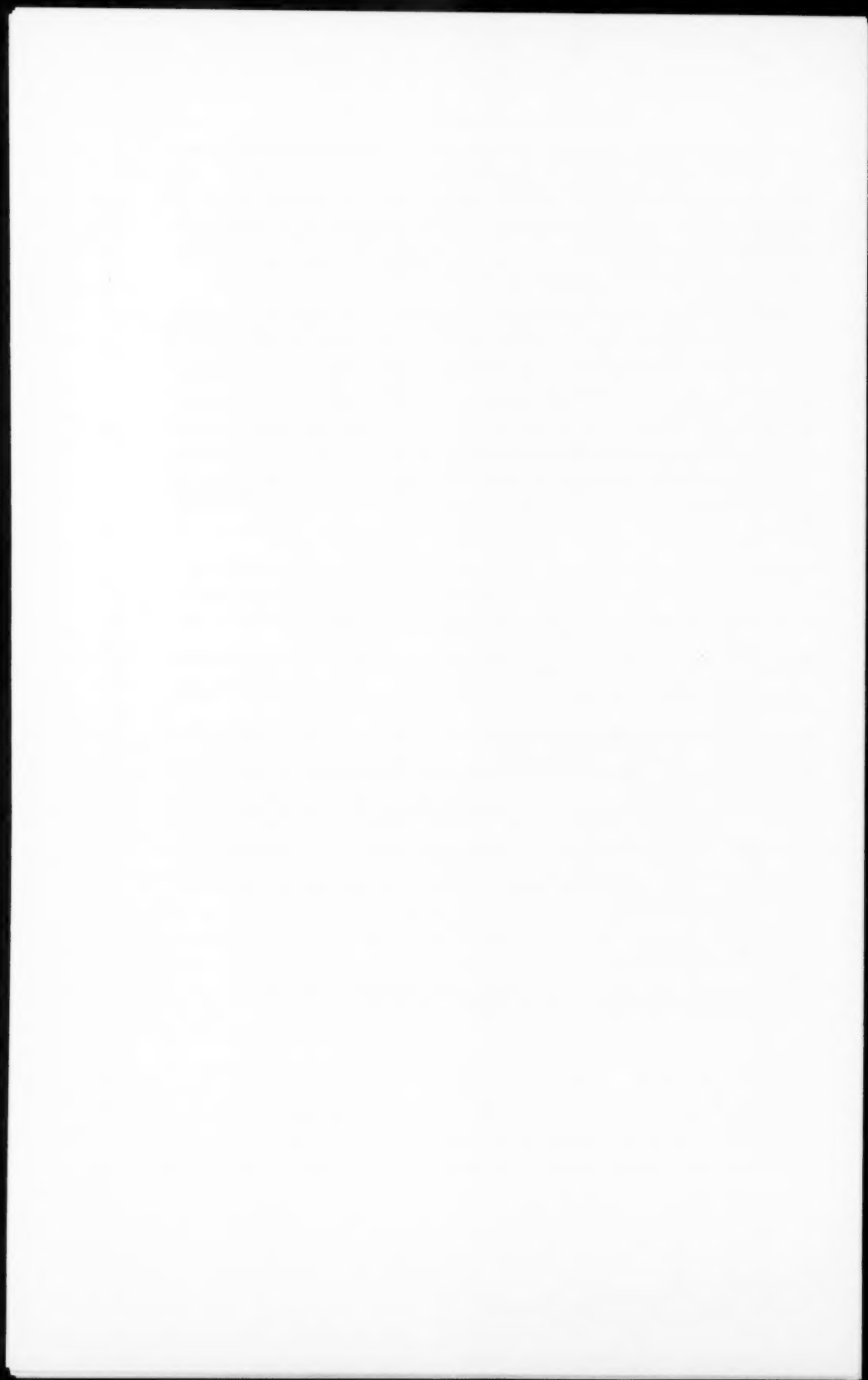
## REFERENCES

1. O. Haupt, *Ein Satz ueber die reellen Raumkurven vierter Ordnung und seine Verallgemeinerung*, Math. Ann., 108 (1933), 126-142.
2. J. Heljelslev, *Introduction à la théorie des suites monotones*, Oversigt Kgl. Danske Vidensk. Selsk. Forh., no. 1 (1914).
3. N. D. Lane, *Differentiable points of arcs in conformal  $n$ -space*, Pac. J. Math., 6 (1956), 301-313.
4. ———, *Characteristic and order of a differentiable point in conformal  $n$ -space*, Trans. Roy. Soc. Can., III, 50 (1956), 47-52.
5. N. D. Lane and P. Scherk, *Differentiable points in the conformal plane*, Can. Jour. Math., 5 (1953), 512-518.
6. ———, *Characteristic and order of differentiable points in the conformal plane*, Trans. Am. Math. Soc., 81 (1956), 358-378.
7. J. Sauter, *Zur Theorie der Bogen  $n$ -ter (Realitäts) Ordnung im projektiven  $R_n$* , Math. Zeits., 42 (1937), 580-592.
8. P. Scherk, *Ueber differenzierbare Kurven und Bögen II*, Časopis pro pěst. mat. a fys., 66 (1937), 172-191.

McMaster University

and the

Summer Research Institute of the Canadian Mathematical Congress



## A Theorem of Friedrichs

HANS ZASSENHAUS, F.R.S.C.

Let  $\mathfrak{A} = \mathfrak{A}(\mathfrak{o}, I)$  be a free associative ring generated by a given set  $I$  of variables over a commutative ring  $\mathfrak{o}$  with unit element. Let  $\mathfrak{A}^* = \mathfrak{A}^*(\mathfrak{o}, I)$  be the associative ring with generators  $\sigma_1 x, \sigma_2 y$  and defining relations

$$(1) \quad \sigma_1 x \cdot \sigma_2 y = \sigma_2 y \cdot \sigma_1 x \quad (x, y \in I)$$

over  $\mathfrak{o}$ . The mapping of  $x$  onto  $\sigma_i x$  of  $I$  into  $\mathfrak{A}^*$  can be uniquely extended to a ring homomorphism  $\sigma_i$  of  $\mathfrak{A}$  into  $\mathfrak{A}^*$ . Also the mapping of  $x$  onto  $\sigma x = \sigma_1 x + \sigma_2 x$  of  $I$  into  $\mathfrak{A}^*$  can be uniquely extended to a ring homomorphism  $\sigma$  of  $\mathfrak{A}$  into  $\mathfrak{A}^*$ .<sup>†</sup>

Friedrichs has stated the following

THEOREM 1. *The set of all elements  $u$  of  $\mathfrak{A}$  for which*

$$(2) \quad \sigma u = \sigma_1 u + \sigma_2 u$$

*is an  $\mathfrak{o}$ -sub-Lie-ring  $L^*$  of  $\mathfrak{A}$  (3; 4).*

Furthermore, Friedrichs has conjectured

THEOREM 2. *If  $\mathfrak{o}$  is a field of characteristic 0 then the  $\mathfrak{o}$ -Lie-ring  $L^*$  defined in theorem 1 is the  $\mathfrak{o}$ -sub-Lie-ring  $L(\mathfrak{o}, I)$  of  $\mathfrak{A}$  that is generated by the members of  $I$  (3; 4).*

This theorem can be generalized as follows. Let, for an arbitrary commutative ring  $\mathfrak{o}$  with unit element,  $L(\mathfrak{o}, I)$  be the  $\mathfrak{o}$ -sub-Lie-ring of  $\mathfrak{A}(\mathfrak{o}, I)$  generated by the members of  $I$ . Let, for any natural number  $n$ ,  $L(\mathfrak{o}, I)^n$  be the submodule of  $\mathfrak{A}(\mathfrak{o}, I)$  generated by the  $n$ th powers of the elements of  $L(\mathfrak{o}, I)$ . For every  $n$  define  $\mathfrak{n}$  to be the sum of  $n$  times the unit element of  $\mathfrak{o}$  and denote by  $\mathfrak{o}_n$  the subring of  $\mathfrak{o}$  that is formed by all elements of  $\mathfrak{o}$  annihilated by  $n$ . Let

$$L(\mathfrak{o}, I)_n = \sum_{i=1}^{\infty} \mathfrak{o}_n L(\mathfrak{o}, I)^{n^i}.$$

We have

THEOREM 3. *The Lie-ring  $L^*$  of Theorem 1 is the direct sum of the  $\mathfrak{o}$ -submodules  $L(\mathfrak{o}, I)$  and  $L(\mathfrak{o}, I)_p$  for all prime numbers  $p$ .*

<sup>†</sup>Indeed  $\sigma_1, \sigma_2, \sigma$  are isomorphisms because  $\tau_i \sigma_i = \tau_i \sigma = 1\mathfrak{A}$  where  $\tau_i$  is that ring homomorphism of  $\mathfrak{A}^*$  onto  $\mathfrak{A}$  over  $\mathfrak{o}$  that maps  $\sigma_i x$  onto  $x$ ,  $\sigma_j x$  onto 0 for  $j \neq i$ ,  $x \in I$ .

An application of this theorem can be made in the case that  $\mathfrak{o}$  is a ring of prime characteristic  $p$ . The characteristic of  $\mathfrak{A}^*$  also is  $p$ . It follows that

$$(3) \quad (A + B)^p = A^p + B^p \text{ if } A, B \in \mathfrak{A}^* \text{ and } AB = BA,$$

as follows from the fact that the binomial coefficients

$$\binom{p}{1}, \binom{p}{2}, \dots, \binom{p}{p-1}$$

are divisible by  $p$  and therefore vanish in  $\mathfrak{o}$ . Since for  $X$  in  $\mathfrak{A}$  we always have  $\sigma_1 X \cdot \sigma_2 X = \sigma_2 X \cdot \sigma_1 X$ , it follows that for  $X$  in  $L^*$

$$\sigma(X^p) = (\sigma_1 X + \sigma_2 X)^p = (\sigma_1 X)^p + (\sigma_2 X)^p = \sigma_1(X^p) + \sigma_2(X^p)$$

and hence  $X^p$  also belongs to  $L^*$ . We have

**THEOREM 4.** *If the characteristic of the coefficient ring  $\mathfrak{o}$  is a prime number  $p$ , then*

- (a)  $L^{*p} \leq L^*$ ,
- (b)  $L^* \circ L^* \leq L(\mathfrak{o}, I)$ ,
- (c)  $(X + Y)^p - X^p - Y^p$  belongs to  $L(\mathfrak{o}, I)$  for any two elements  $X, Y$  of  $L^*$ .

In order to find an application of Theorem 2 we embed  $\mathfrak{A}$  into a power series ring as follows.

Firstly, embed  $\mathfrak{A}$  into the associative ring  $\mathfrak{A}_1$  of all expressions  $\lambda + X$  ( $\lambda \in \mathfrak{o}, X \in \mathfrak{A}$ ) subject to the rules.

$$(4) \quad \begin{aligned} \lambda + X = \mu + Y & \text{ is equivalent to } \lambda = \mu, X = Y \\ (\lambda + X) + (\mu + Y) &= (\lambda + \mu) + (X + Y) \\ (\lambda + X)(\mu + Y) &= \lambda\mu + (\lambda Y + \mu X + XY) \end{aligned}$$

with  $1 = 1 + 0$  as unit element and a basis  $M$  over  $\mathfrak{o}$  that is formed by the monomials  $1$  and

$$x_1 x_2 \dots x_r \quad (r = 1, 2, \dots; x_1, x_2, \dots \text{ contained in } I)$$

of degree 0 resp.  $r = 1, 2, \dots$

Any element  $X \neq 0$  of  $\mathfrak{A}_1$  is a linear combination of finitely many monomials with non-vanishing coefficients in  $\mathfrak{o}$ . The degrees of the monomials contributing to  $X$  have a minimum  $d(X)$ . If we set  $d(0) = +\infty$  then we have defined a non-archimedean Kürschak valuation so that

(5)  $d(X)$  is a non negative integer or  $+\infty$ , uniquely defined for any element  $X$  of  $\mathfrak{A}_1$ ,

$$\begin{aligned} d(X) &= +\infty \text{ if and only if } X = 0 \\ d(XY) &\geq d(X) + d(Y) \\ d(X + Y) &\geq \text{Min}(d(X), d(Y)) \\ d(\lambda X) &\geq d(X) \text{ for } \lambda \text{ contained in } \mathfrak{o}. \end{aligned}$$

The completion  $\mathfrak{P}$  of  $\mathfrak{A}_1$  with respect to  $d$  is called the power series ring of  $I$  over  $\mathfrak{o}$ . The function  $d$  has a unique extension to  $\mathfrak{P}$  such that (5) remains in force. Similarly we embed  $\mathfrak{A}^*$  into the associative ring  $\mathfrak{A}_1^*$  of all expression  $\lambda + X$  ( $\lambda \in \mathfrak{o}, X \in \mathfrak{A}^*$ ) subject to the rules (4), with  $1 = 1 + 0$  as unit element and a basis  $M^*$  over  $\mathfrak{o}$  that is formed by all elements

$$\sigma_1(X_1) \sigma_2(X_2) \quad (X_1, X_2 \in M).$$

We define the degree of the basis element  $X = \sigma_1(X_1)\sigma_2(X_2)$  by the formula  $d^*(X) = d(X_1) + d(X_2)$ . Any element  $Y \neq 0$  of  $\mathfrak{A}^*$  is a linear combination of finitely many basis elements with non-vanishing coefficients in  $\mathfrak{o}$ . The minimum of the degrees of the contributing basis elements may be denoted by  $d^*(Y)$ . We set  $d^*(0) = +\infty$  and a non-archimedean Kürschak valuation  $d^*$  of  $\mathfrak{A}^*$  is thus defined satisfying the rules analogous to (5). It has a unique extension to the completion  $\mathfrak{P}^*$  of  $\mathfrak{A}^*$  with respect to  $d^*$  still satisfying the rules analogous to (5). The isomorphisms  $\sigma_1, \sigma_2, \sigma$  can be uniquely extended to isomorphisms of  $\mathfrak{P}$  into  $\mathfrak{P}^*$  over  $\mathfrak{o}$  such that

$$(6) \quad \sigma_1(1) = \sigma_2(1) = \sigma(1) = 1,$$

$$(7) \quad d^*(\sigma_1 X) = d^*(\sigma_2 X) = d^*(\sigma X) = d(X)$$

for every  $X$  contained in  $\mathfrak{P}$ .

Now we have

**THEOREM 5.** *The set of all elements  $u$  of  $\mathfrak{P}$  satisfying (2) is the completion  $\overline{L}^*$  of  $L^*$  with respect to the Kürschak valuation  $d$ .*

If  $\mathfrak{o}$  is a field of characteristic 0 then for any element  $P$  of the completion  $\overline{\mathfrak{A}}$  with respect to  $d$  or of the completion  $\overline{\mathfrak{A}}^*$  of  $\mathfrak{A}^*$  with respect to  $d^*$  the infinite series

$$(8) \quad \exp(P) = \sum_{i=0}^{\infty} \frac{P^i}{i!}$$

is convergent to an element of  $\mathfrak{P}, \mathfrak{P}^*$  that is congruent to 1 modulo  $\overline{\mathfrak{A}}, \overline{\mathfrak{A}}^*$  respectively. Conversely, for an element  $Q$  of  $\mathfrak{P}, \mathfrak{P}^*$  that is congruent to 1 modulo  $\overline{\mathfrak{A}}, \overline{\mathfrak{A}}^*$  respectively, the infinite series

$$\log Q = \sum_{i=1}^{\infty} (-1)^{i+1} \frac{(Q-1)^i}{i}$$

converges to an element of  $\overline{\mathfrak{A}}, \overline{\mathfrak{A}}^*$  respectively, and there are the identities

$$\log \exp(P) = P, \quad \exp \log(Q) = Q.$$

Moreover

$$(9) \quad \exp(P + P') = \exp(P) \cdot \exp(P') \quad \text{if } PP' = P'P,$$

$$(10) \quad \log(QQ') = \log(Q) + \log(Q') \quad \text{if } QQ' = Q'Q.$$

For any two elements  $X, Y$  of  $\overline{L}^* = \overline{L}(\sigma, I)$  we have

$$\sigma(X) = \sigma_1 X + \sigma_2 X, \quad \sigma(Y) = \sigma_1 Y + \sigma_2 Y$$

and according to the previous identities

$$\begin{aligned}
 \sigma(\log(\exp X \cdot \exp Y)) &= \log(\exp(\sigma X) \cdot \exp(\sigma Y)) \\
 &= \log(\exp(\sigma_1 X + \sigma_2 X) \cdot \exp(\sigma_1 Y + \sigma_2 Y)) \\
 &= \log(\exp(\sigma_1 X) \exp(\sigma_2 X) \cdot \exp(\sigma_1 Y) \exp(\sigma_2 Y)), \\
 &= \log((\exp(\sigma_1 X) \exp(\sigma_1 Y)) \cdot (\exp(\sigma_2 X) \exp(\sigma_2 Y))), \\
 &= \log(\exp(\sigma_1 X) \cdot \exp(\sigma_1 Y) + \log(\exp(\sigma_2 X) \cdot \exp(\sigma_2 Y))) \\
 &= \sigma_1(\log(\exp X \cdot \exp Y)) + \sigma_2(\log(\exp X \cdot \exp Y))
 \end{aligned}$$

so that from Theorem 5 the existence of the Baker-Campbell-Hausdorff formula

$$(11) \quad \log(\exp(X) \cdot \exp(Y)) \in \bar{L}^* \text{ if } X, Y \in \bar{L}^*$$

results. A further consequence will be

**THEOREM 6.** *All elements of the form  $\exp(X)$  with  $X$  contained in  $\bar{L}^*$  form a multiplicative group  $\exp(\bar{L}^*)$  which is the completion of the multiplicative group  $\mathfrak{G}$  that is generated by the elements  $\exp(\lambda x)$  ( $\lambda \in \mathfrak{o}$ ,  $x \in I$ ).*

*Proof of Theorem 1.*  $0$  belongs to  $L^*$ . If  $u, u_1, u_2$  belong to  $L^*$  and if  $\lambda$  belongs to  $\mathfrak{o}$  then

$$\begin{aligned}
 \sigma(\lambda u) &= \lambda \sigma(u) = \lambda \sigma_1(u) + \lambda \sigma_2(u) = \sigma_1(\lambda u) + \sigma_2(\lambda u), \lambda u \in L^*; \\
 \sigma(u_1 + u_2) &= \sigma(u_1) + \sigma(u_2) = \sigma_1(u_1) + \sigma_2(u_1) + \sigma_1(u_2) + \sigma_2(u_2) \\
 &= (\sigma_1(u_1) + \sigma_2(u_1)) + (\sigma_1(u_2) + \sigma_2(u_2)) \\
 &= \sigma(u_1) + \sigma(u_2), \quad u_1 + u_2 \in L^*; \\
 \sigma_1(u_i) \circ \sigma_2(u_k) &= \sigma_1(u_i) \sigma_2(u_k) - \sigma_2(u_i) \sigma_1(u_k) = 0, \\
 \sigma(u_1 \circ u_2) &= \sigma(u_1) \circ \sigma(u_2) = (\sigma_1(u_1) + \sigma_2(u_1)) \circ (\sigma_1(u_2) + \sigma_2(u_2)) \\
 &= \sigma_1(u_1) \circ \sigma_2(u_2) + \sigma_2(u_1) \circ \sigma_2(u_2) + 0 + 0 \\
 &= \sigma_1(u_1 \circ u_2) + \sigma_2(u_1 \circ u_2), \quad u_1 \circ u_2 \in L^*; \text{ as required.}
 \end{aligned}$$

In order to prove Theorems 2,3 we remark that according to (1; 5) the ring  $\mathfrak{A} = \mathfrak{A}(\mathfrak{o}, I)$  is the Birkhoff-Witt embedding ring of  $L(I)$  over  $\mathfrak{o}$ . Moreover  $L(\mathfrak{o}, I)$  is the free Lie-ring of  $I$  over  $\mathfrak{o}$  and it has a homogeneous  $\mathfrak{o}$ -basis.

We generalize Theorems 1,2,3 to

**THEOREM 7.** *Let  $\mathfrak{o}$  be a commutative ring with unit element. Let  $K$  be an  $\mathfrak{o}$ -Lie-ring with an ordered basis  $B$  over  $\mathfrak{o}$ . Let  $\mathfrak{A}$  be the Birkhoff-Witt embedding ring of  $K$  over  $\mathfrak{o}$ . Let  $A^*$  be the  $\mathfrak{o}$ -ring with the generators  $\sigma_1(a), \sigma_2(a)$  ( $a \in A$ ) and the defining relations*

$$\begin{aligned}
 (12) \quad \sigma_i(a + b) &= \sigma_i(a) + \sigma_i(b) \\
 \sigma_i(ab) &= \sigma_i(a) \sigma_i(b) \\
 \sigma_i(\lambda a) &= \lambda \sigma_i(a)
 \end{aligned}$$

for  $i = 1, 2$ ;  $a, b$  contained in  $A$  and  $\lambda$  contained in  $\mathfrak{o}$ ,

$$(13) \quad \sigma_1(a) \sigma_2(b) = \sigma_2(b) \sigma_1(a).$$

Then the mapping of  $L$  into  $A^*$  that maps  $a$  onto  $\sigma_1(a) + \sigma_2(a)$  for every  $a$  of  $K$  can be uniquely extended to an isomorphism  $\sigma$  of  $A$  into  $A^*$  over  $\mathfrak{o}$ .



The set  $K^*$  of all elements  $u$  of  $A$  satisfying (2) is an  $\mathfrak{o}$ -sub-Lie-ring of  $A$  and

$$K^* = K + \sum K_p$$

where  $p$  runs over all natural primes and

$$K_p = \sum_{i=1}^{\infty} \mathfrak{o}_p K^{p^i}.$$

*Proof of Theorem 7.* It follows from (12) that  $\sigma_1, \sigma_2$  are homomorphisms of  $A$  into  $A^*$  over  $\mathfrak{o}$ . The mapping  $\sigma$  of  $L$  into  $A^*$  that maps  $a$  onto  $\sigma(a) = \sigma_1(a) + \sigma_2(a)$  is a homomorphism over  $\mathfrak{o}$  because

$$\begin{aligned} \sigma(a+b) &= \sigma_1(a+b) + \sigma_2(a+b) = \sigma_1 a + \sigma_1 b + \sigma_2 a + \sigma_2 b \\ &= (\sigma_1 a + \sigma_2 a) + (\sigma_1 b + \sigma_2 b) = \sigma a + \sigma b \\ \sigma(\lambda a) &= \sigma_1(\lambda a) + \sigma_2(\lambda a) = \lambda \sigma_1 a + \lambda \sigma_2 a = \lambda(\sigma_1 a + \sigma_2 a) \\ &= \lambda \sigma a \end{aligned}$$

and from (13) it follows that

$$\sigma_1 a \circ \sigma_2 b = \sigma_1 b \circ \sigma_2 a = 0$$

so that

$$\begin{aligned} \sigma(a \circ b) &= \sigma_1(a \circ b) + \sigma_2(a \circ b) = \sigma_1 a \circ \sigma_1 b + \sigma_2 a \circ \sigma_2 b \\ &= (\sigma_1 a + \sigma_2 a) \circ (\sigma_1 b + \sigma_2 b) = \sigma a \circ \sigma b. \end{aligned}$$

Since the Birkhoff-Witt embedding ring  $A$  according to (5) is defined as an associate  $\mathfrak{o}$ -ring containing  $L$  as  $\mathfrak{o}$ -sub-Lie-ring such that  $A$  is generated by  $L$  and such that every  $\mathfrak{o}$ -homomorphism of  $L$  into an associative  $\mathfrak{o}$ -ring can be extended to an  $\mathfrak{o}$ -homomorphism over  $A$  it follows that  $\sigma$  can be uniquely extended to an  $\mathfrak{o}$ -homomorphism of  $A$  into  $A^*$  which also may be denoted by  $\sigma$ . We observe that the mapping  $\tau_i$  that maps  $\sigma_i a$  onto  $a$ , but  $\sigma_k a$  onto 0 for  $a \in A, k \neq i$ , can be uniquely extended to an  $\mathfrak{o}$ -homomorphism of  $A^*$  onto  $A$  because the defining relations of  $A^*$  are preserved. Since

$$\tau_i \sigma_i = \tau_i \sigma = 1_A$$

it follows that  $\sigma_1, \sigma_2, \sigma$  are  $\mathfrak{o}$ -isomorphisms of  $A$  into  $A^*$ .

As in the proof of Theorem 1 it is proved that the set  $K^*$  of all elements  $u$  of  $A$  satisfying (2) is an  $\mathfrak{o}$ -sub-Lie-ring of  $A$ . By definition of  $\sigma$  the  $\mathfrak{o}$ -Lie-ring  $K^*$  contains  $K$  as  $\mathfrak{o}$ -sub-Lie-ring. Also  $K_p$  is contained in  $K^*$  because for  $a \in L, \lambda \in \mathfrak{o}_p, \mu > 0$  we have

$$\begin{aligned} \sigma(\lambda a^{p^\mu}) &= \lambda(\sigma a)^{p^\mu} = \lambda(\sigma_1 a + \sigma_2 a)^{p^\mu} \\ &= \lambda(\sigma_1 a)^{p^\mu} + \lambda(\sigma_2 a)^{p^\mu} + \sum_{i=1}^{p^\mu-1} \binom{p^\mu}{i} \lambda(\sigma_1 a)^i (\sigma_2 a)^{p^\mu-i} \end{aligned}$$

where

$$\begin{aligned} p \mid \binom{p^\mu}{i} \binom{p^\mu}{i} \lambda &= 0 \quad \text{for } i = 1, 2, \dots, p^\mu - 1, \\ \sigma(\lambda a^{p^\mu}) &= \sigma_1(\lambda a^{p^\mu}) + \sigma_2(\lambda a^{p^\mu}), \lambda a^{p^\mu} \in K^*. \end{aligned}$$

According to (1; 5) the elements

$$(14) \quad b_1^{\nu_1} b_2^{\nu_2} \dots b_r^{\nu_r} \\ (r = 1, 2, \dots; \quad b_1, b_2, \dots, b_r \in B; \quad b_1 < b_2 < \dots < b_r; \\ \nu_1 > 0, \nu_2 > 0, \dots, \nu_r > 0)$$

form an  $\mathfrak{o}$ -basis  $C$  of  $A$  over  $\mathfrak{o}$ . From the defining relations (12), (13) of  $A^*$  it follows that the elements

$$(15) \quad \sigma_1(c)\sigma_2(c') \quad (c, c' \in C)$$

form an  $\mathfrak{o}$ -basis  $C^*$  of  $A^*$  over  $\mathfrak{o}$ .

If there is an element  $x$  in  $K^*$  that does not belong to  $K + \sum K_p$ , then let  $x$  be a linear combination of as few a number of elements of  $C$  as possible. We define the weight of each basis element (14) as the sum of the exponents  $\nu_1, \nu_2, \dots, \nu_r$ . Now let (14) be a basis element of maximal weight contributing to  $x$ . If  $r > 1$ , then

$$\begin{aligned} \sigma x - \sigma_1 x - \sigma_2 x &= * + \lambda(\sigma(b_1^{\nu_1} \dots b_r^{\nu_r}) - \sigma_1(b_1^{\nu_1} \dots b_r^{\nu_r}) - \sigma_2(b_1^{\nu_1} \dots b_r^{\nu_r})) \\ &= * + \lambda((\sigma_1 b_1 + \sigma_2 b_1)^{\nu_1} \dots (\sigma_1 b_r + \sigma_2 b_r)^{\nu_r} \\ &\quad - (\sigma_1 b_1)^{\nu_1} \dots (\sigma_1 b_r)^{\nu_r} - (\sigma_2 b_1)^{\nu_1} \dots (\sigma_2 b_r)^{\nu_r}) \\ &= * + \lambda \sigma_1(b_1^{\nu_1} \dots b_{r-1}^{\nu_{r-1}}) \sigma_2(b_r^{\nu_r}) \end{aligned}$$

where  $0 \neq \lambda \in \mathfrak{o}$  and where  $*$  stands for linear combinations of basis elements (15) other than the particular basis element  $\sigma_1(b_1^{\nu_1} \dots b_{r-1}^{\nu_{r-1}}) \sigma_2(b_r^{\nu_r})$ . Since  $\sigma x - \sigma_1 x - \sigma_2 x = 0$ , it follows that  $\lambda = 0$ , which is a contradiction. Hence we have  $r = 1$ ,

$$\begin{aligned} 0 &= \sigma x - \sigma_1 x - \sigma_2 x = * + \lambda(\sigma(b_1^{\nu_1}) - \sigma_1(b_1^{\nu_1}) - \sigma_2(b_1^{\nu_1})) \\ &= * + \lambda((\sigma_1 b + \sigma_2 b_1)^{\nu_1} - (\sigma_1 b_1)^{\nu_1} - (\sigma_2 b_1)^{\nu_1}) \\ &= * + \sum_{i=1}^{\nu_1-1} \binom{\nu_1}{i} \lambda (\sigma_1 b_1)^i (\sigma_2 b_1)^{\nu_1-i} \end{aligned}$$

where again  $\lambda \neq 0$  and  $*$  stands for linear combinations of basis elements (15) other than the particular basis elements

$$(\sigma_2 b_1)^i (\sigma_2 b_1)^{\nu_1-i} \quad (i = 1, 2, \dots, \nu_1 - i).$$

Hence

$$\binom{\nu_1}{i} \lambda = 0 \text{ for } i = 1, 2, \dots, \nu_1 - 1.$$

If  $\nu_1 = 1$  then  $x - \lambda b_1$  is a linear combination of less basis elements (15) than contribute to  $x$ , contained in  $K^*$ , but not in  $K + \sum K_p$ , a contradiction. If  $\nu_1$  is not a prime power, then the greatest common divisor of all binomial coefficients

$$\binom{\nu_1}{i} \quad i = 1, 2, \dots, \nu_1 - 1$$

is 1 and hence  $\lambda = 0$ , a contradiction. If  $\nu_1 = p^a > 1$ ,  $p$  a prime number, then the greatest common divisor of the binomial coefficients

$$\begin{pmatrix} p^\mu \\ 1 \end{pmatrix}, \dots, \begin{pmatrix} p^\mu \\ p^\mu - 1 \end{pmatrix}$$

is  $p$  so that

$$p\lambda = 0, \lambda \in \mathfrak{v}_p, x - \lambda b_1^{p^\mu} \in K^*, x - \lambda b_1^{p^\mu} \notin K + \sum K_p$$

and

$$x - \lambda b_1^{p^\mu}$$

is a linear combination of less basis elements (15) than contribute to  $x$ , again a contradiction. Hence we have  $K^* = K + \sum K_p$  (q.e.d.).

COROLLARY. *It follows from the proof that the elements*

$$(16) \quad b^{p^\mu} (b \in B, \mu > 0)$$

*form a basis of  $K_p$  over  $\mathfrak{o}_p$  and that the sum  $K + \sum K_p$  is a direct sum.*

Generalizing Theorem 4 we have

THEOREM 8. *With the notations of Theorem 7*

$$(17) \quad K^* \circ K^* \leq K$$

$$(18) \quad (\mathfrak{o}_p K + K_p)^p \leq \mathfrak{o}_p K + K_p$$

(19)  $(X + Y)^p - X^p - Y^p \in \mathfrak{o}_p K$  if  $X, Y$  belong to  $\mathfrak{o}_p K + K_p$ ; moreover, the  $\mathfrak{o}$ -Lie-ring  $L^*$  is the normalizer of the  $\mathfrak{o}$ -sub-Lie-ring  $L(\mathfrak{o}, I)$  of  $\mathfrak{A}(\mathfrak{o}, I)$  if  $I$  contains at least two elements.

*Proof of Theorem 8.* The associative ring  $\mathfrak{A}(\mathfrak{o}, I)$  is the direct sum of the  $\mathfrak{o}$ -submodules

$$\mathfrak{A}(\mathfrak{o}, I, x_1^{n_1}, x_2^{n_2}, \dots, x_r^{n_r})$$

formed by the homogeneous polynomials of degree  $n_i$  in  $x_i$  where  $x_1, x_2, \dots, x_r$  are  $r$  distinct elements of  $I$  such that  $x_1 < x_2 < \dots < x_r, r = 1, 2, \dots$  and  $n_1, n_2, \dots, n_r$  are positive. The  $\mathfrak{o}$ -sub-Lie-ring  $L(\mathfrak{o}, I)$  is the direct sum of the  $\mathfrak{o}$ -submodules

$$L(\mathfrak{o}, I) \cap \mathfrak{A}(\mathfrak{o}, I, x_1^{n_1}, x_2^{n_2}, \dots, x_r^{n_r}).$$

Over the ring  $\mathfrak{o}_0$  of the rational integers, each of the latter submodules has a basis

$$B(x_1^{n_1}, x_2^{n_2}, \dots, x_r^{n_r}).$$

It is mapped onto an  $\mathfrak{o}$ -basis

$$B(\mathfrak{o}, x_1^{n_1}, x_2^{n_2}, \dots, x_r^{n_r})$$

of the  $\mathfrak{o}$ -submodule

$$L(\mathfrak{o}, I) \cap \mathfrak{A}(\mathfrak{o}, I, x_1^{n_1}, x_2^{n_2}, \dots, x_r^{n_r})$$

by the homomorphism of  $L(\mathfrak{o}_0, I)$  into  $L(\mathfrak{o}, I)$  that maps  $x$  onto  $x$  for each  $x$  belonging to  $I$ . The union of the basis sets

$$B(\mathfrak{o}, x_1^{n_1}, x_2^{n_2}, \dots, x_r^{n_r})$$

is an  $\mathfrak{o}$ -basis  $B$  of  $L(\mathfrak{o}, I)$ . There is the  $\mathfrak{o}$ -basis  $C$  of  $\mathfrak{A}(\mathfrak{o}, I)$  consisting of the elements (14).

In particular, if  $\mathfrak{o}$  is the field of  $p$  elements ( $p$  a prime number),  $I = \{x, y\}$  then

$$x, y, x + y, x^p, y^p, (x + y)^p$$

belong to  $L^*$  so that the elements  $x^p \circ y$ ,  $(x + y)^p - x^p - y^p$  are linear combinations of the elements

$$b^{\mu} (b \in B; \mu = 0, 1, 2, \dots)$$

over  $\mathfrak{o}$ . But since the Birkhoff ordering operation (1) applied to  $x^p \circ y = x^p y - yx^p$  will remove the contribution made by the basis element  $x^p y$ , there will remain only the contributions made by the elements of  $B(\mathfrak{o}, x^p, y)$ . Similarly it follows that the Birkhoff ordering operation applied to  $(x + y)^p - x^p - y^p$  will remove the contributions made by the basis elements  $x^p, y^p$  so that only the contributions made by the elements of  $B(\mathfrak{o}, x^i, y^{p-i})$  ( $i = 1, 2, \dots, p - 1$ ) will remain. Hence there are identities

$$\begin{aligned} x^p \circ y &= F_1(x, y) + pG_1(x, y) \\ (x + y)^p - x^p - y^p &= \Lambda_1(x, y) + pH_1(x, y) \end{aligned}$$

with  $F_1(x, y)$ ,  $\Lambda_1(x, y)$  belonging to  $L(\mathfrak{o}_0, I)$  and  $G_1(x, y)$ ,  $H_1(x, y)$  belonging to  $\mathfrak{A}(\mathfrak{o}_0, I)$ . Iterating these identities the identities

$$(20) \quad x^{p^\mu} \circ y = F_\mu(x, y) + pG_\mu(x, y)$$

$$(21) \quad (x + y)^{p^\mu} - x^{p^\mu} - y^{p^\mu} = \Lambda_\mu(x, y) + pH_\mu(x, y)$$

with  $F_\mu(x, y)$ ,  $\Lambda_\mu(x, y)$  belonging to  $L(\mathfrak{o}_0, I)$ ,  $G_\mu(x, y)$ ,  $H_\mu(x, y)$  belonging to  $\mathfrak{A}(\mathfrak{o}_0, I)$  for  $\mu = 0, 1, 2, \dots$  are obtained. By application of the identities (20), (21) it follows that

$$K_p \circ K \subseteq \mathfrak{o}_p K, K_p \circ K_p \subseteq \mathfrak{o}_p K$$

and (15). Since for distinct natural prime numbers  $p, q$ , we have

$$0 = pK_p \circ K_q = K_p \circ qK_q$$

and since there are rational integers  $s, t$  satisfying  $ps + qt = 1$  it follows that  $K_p \circ K_q = 0$ . Thus we find (17) and by an application of (21) and (17) we obtain (19).

In particular we have  $L^* \circ L^* \subseteq L(\mathfrak{o}, I)$  for any commutative ring  $\mathfrak{o}$  with unit element. Hence  $L^*$  is contained in the normalizer  $NL(\mathfrak{o}, I)$  of  $L(\mathfrak{o}, I)$  in  $\mathfrak{A}(\mathfrak{o}, I)$ . For any element  $u$  of  $NL(\mathfrak{o}, I)$  and for any element  $a$  of  $L(\mathfrak{o}, I)$  we have  $u \circ a \in L(\mathfrak{o}, I)$ ,

$$\begin{aligned} (\sigma u - \sigma_1 u - \sigma_2 u) \circ \sigma a &= \sigma u \circ \sigma a - \sigma_1 u \circ (\sigma_1 a + \sigma_2 a) - \sigma_2 u \circ (\sigma_1 a + \sigma_2 a) \\ &= \sigma u \circ \sigma a - \sigma_1 u \circ \sigma_1 a - \sigma_2 u \circ \sigma_2 a \\ &= \sigma(u \circ a) - \sigma_1(u \circ a) - \sigma_2(u \circ a) \\ &= 0. \end{aligned}$$

Hence the element  $\sigma u - \sigma_1 u - \sigma_2 u$  is contained in the centralizer of the  $\mathfrak{o}$ -sub-ring  $\sigma \mathfrak{A}(\mathfrak{o}, I)$  of  $\mathfrak{A}^*(\mathfrak{o}, I)$ . We are going to show that the centralizer of  $\sigma \mathfrak{A}(\mathfrak{o}, I)$  in  $\mathfrak{A}^*(\mathfrak{o}, I)$  is 0 if  $I$  consists of more than one element. Then we will conclude that  $\sigma u - \sigma_1 u - \sigma_2 u = 0$ , hence  $u$  belongs to  $L^*$  and  $NL(\mathfrak{o}, O) = L^*$  so that Theorem 8 is established.

Indeed, if there is an element  $z \neq 0$  in the centralizer of  $\sigma \mathfrak{A}(\mathfrak{o}, I)$  in  $\mathfrak{A}^*(\mathfrak{o}, I)$  then it can be represented in the form

$$z = \sum_{k=1}^r \sigma_1(X_k) \sigma_2(c_k)$$

where  $r > 0$  and  $c_1, c_2, \dots, c_r$  are distinct members of the  $\mathfrak{o}$ -basis  $C$  of  $\mathfrak{A}(\mathfrak{o}, I)$  over  $\mathfrak{o}$  formed by the elements (14) such that the degree of the homogenous polynomials  $c_1, c_2, \dots, c_r$  form a never decreasing finite sequence and, moreover,  $0 \neq X_k \in \mathfrak{A}_1(\mathfrak{o}, I)$  for  $k = 1, 2, \dots, r$ . For any element  $x$  of  $I$  we find that

$$\begin{aligned} 0 &= \sigma x \circ z = \sigma_1 x \circ z + \sigma_2 x \circ z \\ &= \sum_{k=1}^r (\sigma_1 x \circ \sigma_1(X_k)) \sigma_2(c_k) + \sum_{k=1}^r (\sigma_1(X_k)) \sigma_2(x \circ c_k) \\ &= \sigma_1(x \circ X_1) \sigma_2(c_1) + \sum_{k=2}^r \sigma_1(Y_k) \sigma_2(c_k), \end{aligned}$$

where  $Y_k \in \mathfrak{A}_1(\mathfrak{o}, I)$ ,  $r \leq s$ , and the elements  $c_{r+1}, \dots, c_s$  are members of  $C$  distinct among themselves and distinct from  $c_1, c_2, \dots, c_r$  such that the degree of each of the elements  $c_{r+1}, \dots, c_s$  is greater than the degree of  $c_1$ . This is because the degree of  $x \circ c_k$  is greater than the degree of  $c_k$  and, *a fortiori*, greater than the degree of  $c_1$ . Since the elements (15) form an  $\mathfrak{o}$ -basis of  $\mathfrak{A}^*(\mathfrak{o}, I)$  it follows that  $x \circ X_1 = 0$  for all members  $x$  of  $I$  so that  $X$  belongs to the centre of  $\mathfrak{A}(\mathfrak{o}, I)$ . The centre of  $\mathfrak{A}(\mathfrak{o}, I)$  is 0 if  $I$  contains at least two elements. In this case, then,  $X_1 = 0$ , a contradiction. Therefore the centralizer of  $\sigma \mathfrak{A}$  in  $\mathfrak{A}^*$  vanishes in the case that  $I$  contains at least two elements (q.e.d.).

*Proof of Theorem 6.* Since the exponential function is bi-continuous with respect to  $d$ , it follows that  $\exp(L^*)$  is closed. The mapping of  $\lambda$  onto  $\exp(\lambda x)$  is an isomorphism of the additive group of  $\mathfrak{o}$  onto a multiplicative group  $\mathfrak{g}(x)$  for each  $x$  of  $I$ . If for  $r$  elements  $x_1, x_2, \dots, x_r$  of  $I$  we have  $x_1 \neq x_2, \dots, x_{r-1} \neq x_r$  and if the elements  $x_1, x_2, \dots, x_r$  of  $\mathfrak{o}$  do not vanish then the power series development of

$$\exp(\lambda_1 x_1) \exp(\lambda_2 x_2) \dots \exp(\lambda_r x_r)$$

has the term  $\lambda_1 \lambda_2 \dots \lambda_r x_1 x_2 \dots x_r$  so that the power series is not 1. Hence the group  $\mathfrak{G}$  is a free product of the generating subgroups  $\mathfrak{g}(x)$  ( $x \in I$ ).

All elements  $X$  of  $\mathfrak{G}$  for which  $d(X - 1) > n$ , form a normal subgroup  $\mathfrak{G}_n$ , consisting of elements of the form

$$X = 1 + \theta_n(X) + \rho_n(X)$$

where  $\theta_n(X)$  belongs to  $\mathfrak{A}_n(\mathfrak{o}, I)$  and  $d(\rho_n(X)) > n$ . By application of the Baker-Hausdorff formula we find that  $\mathfrak{G}$  is contained in  $\exp(\overline{L^*})$ , hence

$$\begin{aligned}\overline{L^*} &\rightarrow \log X \equiv \theta_n(X) \pmod{\mathfrak{A}^{n+1}(\mathfrak{o}, I)} \\ \theta_n(X) &\in L^* \cap \mathfrak{A}_n(\mathfrak{o}, I).\end{aligned}$$

It is clear that  $\theta_n$  is a homomorphism of  $\mathfrak{G}_n$  into the module  $L^* \cap \mathfrak{A}_n(\mathfrak{o}, I)$  with  $\mathfrak{G}_{n+1}$  as kernel. We also observe that the two congruences

$$X \equiv 1(\mathfrak{A}^n), Y \equiv 1(\mathfrak{A}^m) \quad (n > 0, m > 0)$$

imply

$$\begin{aligned}XY - (1 + X \circ Y)YX &= XY - YX - (X \circ Y)YX = -X \circ Y(YX - 1) \\ &= -((X - 1) \circ (Y - 1))(Y(X - 1) + (Y - 1)) \\ &\equiv 0(\mathfrak{A}^{n+m+1}), \\ XYX^{-1}Y^{-1} &\equiv 1 + (X - 1) \circ (Y - 1)(\mathfrak{A}^{n+m+1}).\end{aligned}$$

Since  $\theta_1\mathfrak{G}_1 = L^* \cap \mathfrak{A}_1(\mathfrak{o}, I)$ , it follows that  $\theta_{n+1}(\mathfrak{G}_{n+1}) \supseteq \theta_1\mathfrak{G}_1 \circ \theta_n\mathfrak{G}_n = (L^* \cap \mathfrak{A}_1(\mathfrak{o}, I)) \supseteq \theta_n\mathfrak{G}_n$  and from

$$(L^* \cap \mathfrak{A}_1(\mathfrak{o}, I)) \circ (L^* \cap \mathfrak{A}_n(\mathfrak{o}, I)) = L^* \cap \mathfrak{A}_{n+1}(\mathfrak{o}, I)$$

it follows now by induction over  $n$  that

$$\theta_n(\mathfrak{G}_n) = L^* \cap \mathfrak{A}_n(\mathfrak{o}, I).$$

For any element  $X$  of  $\exp(\overline{L^*})$  distinct from 1 we have

$$0 < d(X - 1) = n < +\infty, X - 1 \equiv \log X(\mathfrak{A}^{n+1}(\mathfrak{o}, I)), \\ \log X \in \overline{L^*},$$

hence there exists an element  $g(X)$  in  $\mathfrak{G}_n$  such that  $g(X) \equiv X(\mathfrak{A}^{n+1}(\mathfrak{o}, I))$ . By induction over  $n$  and by application of the Baker-Hausdorff formula we find that  $X = \lim X_m X_{m-1} \dots X$  where

$$X_1 = g(X), X_m X_{m-1} \dots X_1 \in \mathfrak{G}, XX_1^{-1} \dots X_m^{-1} \in \exp(\overline{L^*}), X_{m+1} = g(XX_1^{-1}X_2^{-1} \dots X_m^{-1}).$$

Hence  $X \in \overline{\mathfrak{G}}$ .

#### REFERENCES

1. G. Birkhoff, *Representation of Lie algebras and Lie groups by matrices*, Ann. of Math., 38 (1937), 526-532.
2. Paul Moritz Cohn, *Sur le critère de Friedrichs pour les commutateurs dans une algèbre associative libre*, C.R. Acad. Sci. Paris, 239 (1954), 743-745.
3. K. O. Friedrichs, *Comm. Pure Appl. Math.*, 6 (1953), 1-72.
4. Wilhelm Magnus, *On the exponential solution of differential equations for a linear operator*, *Comm. Pure Appl. Math.*, 7 (1954), 649-673.
5. Ernst Witt, *Treue Darstellung Liescher Ringe*, J. Reine Angew. Math., 177 (1937), 152-160.

McGill University

and the

Summer Research Institute of the Canadian Mathematical Congress

PRINTED IN CANADA  
AT THE UNIVERSITY OF TORONTO PRESS



